

A STUDY OF IMPLICIT LEARNING
THROUGH AUDITORY STIMULI

by

Asli Aslan

A DISSERTATION

Presented to the Faculty of
The Graduate College at the University of Nebraska
In Partial Fulfillment of Requirements
For the Degree of Doctor of Philosophy

Major: Psychology

Under the Supervision of Professor John H. Flowers

Lincoln, Nebraska

December, 2003

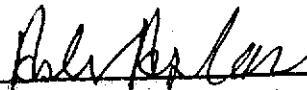
Dissertation Title:

A Study of Implicit
Learning through
Auditory Stimuli

Committee Members: John H. Flowers, Ph.D.
(Committee Chair)
Daniel J. Bernstein, Ph.D.
Calvin P. Garbin, Ph.D.
Cynthia Willis Esqueda
David Moshman, Ph.D.

Program Area: Cognitive

Graduation Date: December 20, 2003



By: Asli Aslan, Ph.D.

DISSERTATION TITLE

A Study of Implicit Learning

Through Auditory Stimuli

BY

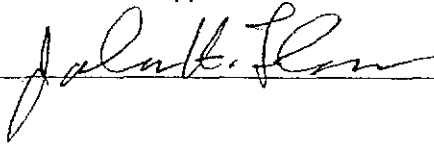
Asli Aslan

SUPERVISORY COMMITTEE:

Approved

Date

Signature



11/12/03

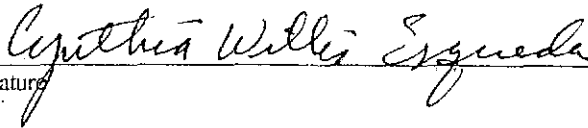
Typed Name

David Washman

Signature

11/12/03

Typed Name



Signature

11-12-03

Typed Name



Signature

11-12-03

Typed Name



Signature

11-12-2003

Typed Name

Signature

Typed Name

A STUDY OF IMPLICIT LEARNING THROUGH AUDITORY STIMULI

Asli Aslan, Ph. D.

University of Nebraska, 2003

Adviser: John H. Flowers

Implicit learning can be defined as unconscious understanding of the relationships between the elements of a knowledge system that occurs without an explicit intention toward discovering the underlying structure of those relationships. Researchers studying implicit learning have constructed a wide variety of knowledge systems most of which are based on probabilistic relationships between visual cues and tried to determine how well the participants of experiment learned those relationships without showing an explicit awareness of learning. However, very few researchers introduced experimental tasks based on learning probabilistic relationships between the structural elements of music, such as tone, key and tempo. The current research ventures into the unexplored field of implicit learning in auditory perception of melodies and tries to discover the appropriate combinations of musical elements whose relationships can be learned implicitly. Musical melodies are employed as stimuli in a simple memory task and attempts are made to determine whether or not participants performed the task better by learning the probabilistic context of the clues embedded in the melodies. The results indicate that the occurrence of implicit learning in processing of melodies is strongly tied to the completeness of the melodies as auditory patterns.

TABLE OF CONTENTS

I.	Definitions of Implicit Learning	1
II.	Fundamental Characteristics of Implicit Learning	3
	A. Absence of Intention or Awareness	3
	B. Absence of Explicitly Stated Rules	4
	C. Absence of Explicitly Presentable Conclusions	5
III.	Implicit Learning Demonstrated in Performing Tasks	6
	A. Prominence of Implicit Learning in Complex Tasks	6
	B. Types of Tasks Utilized in Implicit Learning Studies	8
	1. Artificial Grammar Tasks	8
	2. Control Tasks	10
	3. Sequence Learning Tasks	11
	C. Experimental Procedures Utilized in Implicit Learning Studies	13
	1. Briefing Phase	13
	2. Experimental Phase	14
	3. Testing Phase	15
IV.	A Background on the Structure of Music	17
	A. Music as a Complex Structure of Sound and Tones	17
	B. Music as a Form of Grammar	20
	C. Music as a Collection of Patterns	22

V.	Implicit Learning in Perception of Music.....	23
A.	Pattern Recognition in Music Perception Studies.....	24
B.	Expectancy in Perception of Music Perception Studies.....	25
C.	Lack of Studies on Implicit Learning in Music Perception.....	26
VI.	A General Description of the Experiments.....	28
A.	The Purpose of the Experiments.....	28
B.	General Method of the Experiments.....	29
1.	A Description of the Experimental Task.....	29
2.	The Method of Administration.....	30
3.	Experimental Stimuli.....	31
4.	The Follow-up Questionnaire.....	31
C.	The Predictive Context of the Experiments.....	32
1.	The Purpose of Clues.....	32
2.	Predictiveness Relationships of Clues to Same-Trials.....	33
3.	Melody Clues.....	34
4.	Rhythm Clues.....	35
D.	Reversal and Nonreversal groups.....	36
1.	Reversal of the Predictiveness Relationships.....	36
2.	Differences between Reversal and Nonreversal Groups.....	37
E.	An Overview of Experimental Data.....	39
1.	Independent Variables of the Experiments.....	39

a)	Type of the Clue.....	39
b)	Playback Tempo	39
2.	Performance Measures (Dependent Variables) Targeted by Analysis.....	40
a)	Accuracy Score	40
b)	Reaction Time	40
F.	An Overview of Data Analysis Methods.....	41
1.	Effects of Implicit Learning on Performance.....	41
2.	Analysis of Performance Changes before the 90 th Trial	42
3.	Analysis of Performance Changes after the 90 th Trial	43
4.	Expected Performance Changes With Respect to Clues	44
5.	Statistical Tests Utilized in Analyses	45
VII.	Overview of Experiment 1	47
A.	Purpose	47
B.	Method.....	47
C.	Participants	48
VIII.	Results of Experiment 1	48
A.	Analysis of Performance Changes before the 90 th Trial	48
1.	Mean Scores of Same-Trials.....	48
2.	Mean Scores of Different-Trials.....	51
3.	Mean Scores of Positive-Clue Trials.....	54
4.	Mean Scores of Negative-Clue Trials	55

5.	Analysis of Performance Changes after the 90 th Trial	56
B.	Summary of the Results	56
IX.	Overview of Experiment 2	58
A.	Purpose	58
B.	Method.....	58
C.	Participants	59
X.	Results of Experiment 2	60
A.	Analysis of Performance Changes before the 90 th Trial	60
1.	Mean Scores of Same-Trials.....	60
2.	Mean Scores of Different-Trials.....	63
3.	Mean Scores of Positive-Clue Trials.....	66
4.	Mean Scores of Negative-Clue Trials	67
B.	Analysis of Performance Changes after the 90 th Trial	68
1.	Score Drops In the Reversal and Nonreversal groups	68
2.	Mean Scores of Trials after 90	70
a)	Mean Scores of Same-Trials in the Reversal group	70
b)	Mean Scores of Different-Trials in the Reversal group.....	72
c)	Mean Scores of Same-Trials in the Nonreversal group.....	75
d)	Mean Scores of Different-Trials in the Nonreversal group.....	77
C.	Summary of the Results	79
XI.	Overview of Experiment 3	81

A.	Purpose	81
B.	Method.....	81
C.	Participants	82
XII.	Results of Experiment 3	83
A.	Analysis of Performance Changes before the 90 th Trial	83
1.	Mean Scores of Same-Trials.....	83
2.	Mean Scores of Different-Trials.....	86
3.	Mean Scores of Positive-Clue Trials.....	89
4.	Mean Scores of Negative-Clue Trials	90
B.	Analysis of Performance Changes after the 90 th Trial	91
1.	Score Drops in the Reversal and Nonreversal groups	91
2.	Mean Scores of Trials after 90	93
a)	Mean Scores of Same-Trials in the Reversal group	93
b)	Mean Scores of Different-Trials in the Reversal group.....	95
c)	Mean Scores of Same-Trials in the Nonreversal group.....	98
d)	Mean Scores of Different-Trials in the Nonreversal group.....	101
C.	Summary of the Results	103
XIII.	Overview of Experiment 4	105
A.	Purpose	105
B.	Method.....	105
C.	Participants	106

XIV. Results of Experiment 4	107
A. Analysis of Performance Changes before the 90 th Trial	107
1. Mean Scores of Same-Trials.....	107
2. Mean Scores of Different-Trials.....	109
3. Mean Scores of Positive-Clue Trials.....	112
4. Mean Scores of Negative-Clue Trials	113
B. Analysis of Performance Changes after the 90 th Trial	114
1. Score Drops In the Experiment and Nonreversal groups	114
2. Mean Scores of Trials after 90	115
a) Mean Scores of Same-Trials in the Reversal group	115
b) Mean Scores of Different-Trials in the Reversal group.....	117
c) Mean Scores of Same-Trials in the Nonreversal group.....	119
d) Mean Scores of Different-Trials in the Nonreversal group.....	121
C. Summary of the Results	123
XV. Discussions on the Results of the Experiments	125
A. A Discussion on the Results of Experiment 1	125
B. A Discussion on the Results of Experiment 2	126
C. A Discussion on the Results of Experiment 3	127
D. A Discussion on the Results of Experiment 4.....	128
E. A Discussion on the Responses to Questionnaires.....	129
XVI. Conclusions.....	130

References	134
Appendix A: Analysis of the Task Performance in Experiment 1 after the 90 th Trial	140
1. Score Drops In the Experiment and Nonreversal groups	140
2. Mean Scores of Trials after 90	141
a) Mean Scores of Same-Trials in the Reversal group	141
b) Mean Scores of Different-Trials in the Reversal group.....	143
c) Mean Scores of Same-Trials in the Nonreversal group.....	145
d) Mean Scores of Different-Trials in the Nonreversal group	147
Appendix B: Mean Reaction Times of Experiment 1.....	149
Appendix C: Mean Reaction Times of Experiment 2.....	153
Appendix D: Mean Reaction Times of Experiment 3	157
Appendix E: Mean Reaction Times of Experiment 4.....	161
Appendix F: The Post-Task Questionnaire	165

Chapter 1

A BACKGROUND ON IMPLICIT LEARNING

I. DEFINITIONS OF IMPLICIT LEARNING

Reber (1993) lists several articles from middle and late 1960's as the earliest sources where the term "implicit learning" is mentioned. However, as we understand from that article by Reber, the concept of implicit learning has been mentioned by even earlier sources under different names, such as "learning without awareness" and "incidental learning." The common emphasis of those early terms was the lack of intention to learn or the lack of awareness during learning, because those were the first observed characteristics that were not traditionally associated with the concept of learning. Another distinctive characteristic of this new form of learning was the inability to express the acquired knowledge explicitly. In the absence of precisely defined mechanisms explaining how implicit learning occurred, definitions of implicit learning mostly described its fundamental characteristics.

Attempts to define implicit learning through its fundamental characteristics have sometimes led to overly simplified and somewhat circular definitions. Lewicki (1986), for example, defined implicit learning as the direct opposite of "explicit learning," the usual form of learning that leads to explicitly retrievable knowledge. This was a limited definition focusing solely on the outcome of implicit learning. Underwood and Bright (1996) labeled this approach as an extreme position compared to the rest of the literature on

implicit learning. Reber (1993), for example, warns specifically against “the polarity fallacy” that he describes as viewing implicit and explicit learning as opposing processes. In fact, he does not even favor viewing them as separate and independent processes.

Mainstream definitions encompass other distinctive characteristics of implicit learning. Reber (1976, 1989), for example, defined implicit learning as an “unconscious abstraction process” that gave rise to “abstract knowledge.” Cleeremans and Jimenez (1996) pointed to the acquisition of knowledge without intention as the definitive characteristic of implicit learning. These two authors also noted that this unconscious form of learning would implicitly influence behavior. A definition offered by Perruchet and Vinter (1998) also specifically mentioned the influence of implicit learning on behavior. They suggested that behavior could be affected by “structural features of an experienced situation” even if experiences were not gained through “intentional exploitation of ... explicit knowledge.”

For the purpose of summarizing the basic hypotheses underlying this dissertation, implicit learning can be defined as unconscious understanding of the relationships between the elements of a knowledge system that occurs without an explicit intention toward discovering the underlying structure of those relationships. This choice of definition will help readers to understand the reasons behind the methodology of the experiments that formed the basis for this dissertation. In order to reach a better understanding of those experimental methods, it is now necessary to briefly describe the fundamental characteristics of implicit learning as described by theorists and observed by

experimentalists. The next step will be a review of the types of tasks and analysis methods utilized in previous studies of implicit learning.

II. FUNDAMENTAL CHARACTERISTICS OF IMPLICIT LEARNING

A. Absence of Intention or Awareness

As it was noted in the previous section, implicit learning was first described as a form of learning that occurred without intention or awareness. However, both the absence of intention and the absence of awareness were considered to be claims contradicting the nature of learning. Even before the term "implicit learning" was coined, Postman (1964) noted that accepting that learning could occur without intention was equivalent to "accepting the null hypothesis." He warned that researchers should not take the absence of intention in implicit learning in an absolute sense. Likewise, the absence of awareness should not be taken in any literal sense, either. Frensch (1998) noted that the equivalence of implicit learning to learning without awareness should be taken to mean that "learner is not aware that learning occurs." He also made an excellent point by reminding that there was no scientifically acceptable method of measuring awareness. Thus it is not possible to define a measurable level of awareness below which a learning process will be labeled as implicit learning. The mere presence or absence of awareness is the key to determining whether or not a certain learning process is implicit.

In summary, one of the fundamental characteristics of implicit learning is that, the learners acquire a subconscious understanding of a knowledge system without intentionally making an effort to uncover the underlying rules of the system. However, in order to

compel the learners to acquire the knowledge that they do not target, they must be presented with tasks that will require that hidden knowledge for a successful performance.

B. Absence of Explicitly Stated Rules

As explained above, either the absence of intention to learn or the absence of awareness can be employed as evidences that implicit learning occurs, but only if the knowledge acquisition is indirect. A common method of ensuring that is to ask the participants of a study to perform a task in which they will succeed only if they make use of some rules that are not explicitly stated. For example, in the artificial grammar tasks of Reber (1967, 1969, 1976) and in the control of a sugar production simulation devised by Berry and Broadbent (1984) participants were claimed to be making use of the rules hidden from them. The participants had to have acquired the knowledge of the rules through implicit learning, because they could not verbally explain what those rules were.

It should be noted that, in none of those experiments did the participants make a conscious effort to determine whether or not some rules existed or what those rules were. In fact, as Mathews et. al. (1989) reminded in their literature review, conscious attempts to discover the underlying rules do not improve the performance and explicit instructions may even impair the performance. They refer to Reber (1989) and Berry and Broadbent (1989) who compared the performances of participants who did or did not get explicit instructions relating to the rules of the tasks. For instance, in a study, Flowers and Green (1991) compared the influences of explicit instruction and uninstructed implicit learning of correlations among visual events, in a fine motor task. They also suggested that instructed

and noninstructed participants showed different patterns of motor activity. Their findings suggested a qualitatively different approach to visual information processing and task performance for the instructional conditions. Providing explicit probability instructions was shown to negatively influence task performance.

In summary, another fundamental characteristic of implicit learning is that, the learners are not given any explicit information about the underlying rules that they are expected to discover. In fact, they should not even be explicitly asked to discover those rules, but they must be presented with tasks in which successful performance does not require an explicit understanding of the rules.

C. Absence of Explicitly Presentable Conclusions

Among the primary aspects of implicit learning, Berry and Dienes (1993) pointed to “low rates of transfer to free recall” and to “limited transfer to related tasks.” In other words, participants of the experiments were unable to verbally explain or make use of the knowledge that they acquired implicitly. These findings implied dissociation between verbalizable knowledge and successful performance in experimental tasks. However, they also noted that such conclusions might have been due to “methodological inadequacies.” As explained by Sanderson (1989), the conclusions of many implicit learning studies relied on post-task questionnaires to assess verbalizable knowledge, but questionnaires could not serve as measuring tools. Nevertheless, overwhelming evidence from questionnaires leads researchers to believe that the absence of verbalizable conclusion is a direct result of implicit learning. As Reber (1993) noted, even members of a psychology department could

not clearly explain the patterns of stimuli after successfully performing a task designed by Lewicki. Their extensive knowledge of the subject and of the task should have given them the ability to summarize any verbalizable knowledge (Lewicki et al., 1988)

In summary, this last characteristic of implicit learning is not a prerequisite for the occurrence of implicit learning in experimental tasks, but it can provide the researchers with data proving the presence of implicit learning. The view taken in this study is that, participants themselves may not be able to reliably conclude whether they learnt something expected or unexpected during any learning task. Therefore, it is preferable to draw conclusions from the changes in participants' performances on the given tasks.

III. IMPLICIT LEARNING DEMONSTRATED IN PERFORMING TASKS

As Reber (1993) noted, people tend to first seek to obtain information about the structure underlying their environment and then to utilize that information about the structure. The interesting thing is that, they seem to succeed in doing that even if the underlying structure is of a complex and nonobvious nature. In fact, implicit learning seems to be especially relevant to the performance of complex tasks. Therefore, this part of the discussion will expand upon the relevance of implicit learning to the complexity of tasks.

A. Prominence of Implicit Learning in Complex Tasks

Reber (1993) provides an extensive summary of studies where participants acquired and utilized information about highly complex task structures. In artificial grammar experiments of Reber (e.g., Reber, 1967) participants memorized long strings of letters

following certain artificial grammar rules and then successfully distinguished letter strings that did or did not conform to those rules. Reber and Milward (1971) found that participants could anticipate certain events even if the probabilities of occurrence for those events changed imperceptibly slowly. The participants' ability to exploit the structure hidden in a series of trials was found to be beyond "the limits on explicit recall." Complex rules of production and social systems simulations were also successfully exploited by participants in experiments involving control tasks. Despite of the small random fluctuations in the inter-dependencies between the parameters of such tasks, participants responded appropriately in situations that required them to achieve the desired levels of target parameters.

Conclusions drawn from such experiments commonly state that implicit learning seems to be the primary mode of learning in experiments involving complex tasks. However, the complexity of the task is neither a prerequisite nor a facilitating factor for implicit learning. It is rather a limiting factor that will rule out explicit learning. Reber (1993) makes that clear by stating that implicit learning will not be observed in cases where participants can discover the hidden rules through "simple testing of consciously held hypotheses." Complexity of underlying structure also makes it more likely that participants will succeed at a task only through utilizing the information hidden in the structure by the experimenters. Furthermore, complexity and specificity of hidden information will help the researchers easily determine whether the participants gained an explicit awareness of the underlying structure.

B. Types of Tasks Utilized in Implicit Learning Studies

The above discussion makes it clear that participants of experiments are able to acquire knowledge about very complex sets of rules through implicit learning. The following brief discussion explains what types of tasks and what types of rule systems were employed in implicit learning studies. Some conclusions are drawn from the results of those tasks to show the relevance of implicit learning studies to everyday life and to the study presented in this paper.

1. Artificial Grammar Tasks

The foundation for such studies has been laid by a study of Chomsky and Miller (1958) on finite-state grammars. A finite-state grammar structure is a "Markovian system" that is basically a set of connections between nodes of a graph (a mesh structure) that represent the elements of the grammatical structure (Underwood and Bright, 1996). The nodes connected to each other represent the pairs of grammar elements that can follow each other. The number of nodes and the multiplicity of connections in such a Markovian system determine the level of complexity of the system.

Since grammatical structures can easily be constructed at any level of complexity, artificial grammar studies have long been one of the richest domains for the studies of implicit learning. In fact, some of the earliest studies of implicit learning employed grammar learning tasks (e.g., Reber, 1965). As Reber (1993) states, there are two essential components in grammar-learning studies: the "acquisition phase" and the "testing phase." During the first phase, participants are expected to acquire an understanding of the grammatical rules that are used to construct the sets of letter strings given to them. In the

second phase, the level of their understanding is determined. Participants are given no hints about the existence of any underlying rules until the testing phase (Manza et. al. 1998). The testing phase typically consists of presenting the participants with new strings that may or may not be constructed according to the same grammar rules to which they have been exposed. As Hoffman and Koch (1998) stated, the role of implicit learning in artificial grammar studies is evident in participants' success at accurate classification of test strings (success rates reach 70%). Another evidence is the participants' inability to explain the rules that they utilized. The results from such studies have been used to construct computational models of implicit learning such as the classifier systems mentioned by Dienes (1993).

An interesting side of artificial grammar tasks is that, the participants' performances are found to be independent of what elements are used to form the grammar system. For example, Reber (1969) presented each one of his participants with two sets of letter strings. He found that a typical subject's performance was adversely affected when the two sets of strings were constructed with different sets of rules. However, he found no significant performance change when the two sets of strings used the same rules, even when the letters used in the two sets were different. This finding is especially relevant to the study presented in this paper, because it is also a type of grammar learning study where musical melodies substitute for letter strings. The independence of learning performance from the types of elements leads to an expectation that implementing a single set of rules in different melodies will not have detrimental effects on participants' performances.

2. Control Tasks

Berry (1993) reports that the interest in this type of tasks started in 1970s and refers to two early studies based on simulations of transportation and economic systems (Broadbent, 1977; Broadbent and Aston, 1978). In those studies, the participants' success at correctly manipulating the input parameters was found to be improving as they progressed in their tasks. However, their ability to explain their actions did not improve along with their improving performance. Neither did the researchers find a correlation between the level of success of a subject and the ability to verbalize the relationships utilized by the subject (Broadbent, 1977).

A series of experiments performed by Berry and Broadbent (1984, 1987, 1988) based on simulations of manufacturing operations and social interactions of human beings are the most frequently quoted examples of control tasks. Production control tasks required the participants to adjust certain parameters like workers, wages, and production levels to achieve the targeted levels of productivity or labor peace. As it was summarized by Reber (1993), the target parameters' values depended on the values of input parameters through some multi-variable linear relationships, but the presence of random fluctuation terms ensured that the interrelationships between the input and output parameters would not be easily discovered. According to the researchers' conclusions, despite of the complexity of such tasks, participants have been able to maintain the values of the target parameters within very narrow ranges around the desired values.

The results obtained from the control studies based on such realistic tasks imply that implicit learning is crucial to the performance of any task involving the control of a

complex system without relying on specific rules. Thus it is clear that implicit learning is relevant to everyday lives of people and the jobs that they perform. The dissociation between successful performance and verbally explicable rules leads to the expectation that the job performance can be manipulated through implicit understanding of rules embedded in visual and auditory clues embedded in the work environment. The fact that explicit instructions did not help improve the task performance (Berry and Broadbent, 1984) implies that it may be preferable to make job instructions implicit rather than explicit.

3. Sequence Learning Tasks

Cleeremans and Jimenez (1998) note that sequence learning paradigm has recently been favored by more and more researchers attempting to answer some fundamental questions about the nature of implicit learning. Shanks and Johnstone (1998) claimed that sequence learning tasks are “particularly well-suited to studying implicit learning.”

A typical sequence learning task asks the participants to respond to events that will have some predictable characteristics, such as color, sound, the location of occurrence, etc. In each trial, the target characteristic takes on one of the alternative values or forms and follows a sequence pre-determined by the researchers. The sequence is constructed according to certain rules of probability that the participants are expected to learn and exploit while performing the task. A decreasing trend in response times (or, reaction times) of a subject is taken to be an evidence that the subject learned the sequence. However, there is no standard method of determining whether or not the learning was implicit or explicit. One method is to ask the subject predict the next event in the sequence, like it was done by

Cleeremans and McClelland (1991). They asked their subject to predict the location of the next stimulus instead of merely responding to the present one. However, responses of the participants to such inquiries failed to distinguish the conscious learning from unconscious learning (Cleeremans and Jimenez, 1998).

According to the results of some experiments, the complexity of the rules underlying a sequence learning task will help eliminate any contribution from explicit learning. By briefing their participants about the simple rules used in a sequence learning task, Frensch and Miner (1996) showed that a typical subject's success was positively affected by explicit instructions. However, explicit instructions gave no advantages to participants performing complex tasks (Jimenez et al., 1996). The dissociation between explicit instructions and the successful task performance implies the prominence of implicit learning in complex sequence learning tasks. Introducing random noise or probabilistic rules to sequences is another way to reduce the participants' tendency to obtain conscious knowledge of the underlying rules (Cleeremans and McClelland, 1991). However, according to Cleeremans and Jimenez (1998) "deterministic sequences still tend to elicit more learning than probabilistic sequences." The two authors also argue that randomly inserting sequence elements to otherwise-fixed sequences disrupts participants' learning process and makes the responses to predictable events as bad as the responses to unpredictable ones. This last argument has a special relevance to the study presented in this paper, because it implies that the responses of participants to non-deterministic clues must be analyzed with great care. Probabilistic nature of the predictive context of the present

study is certain to make it difficult to evaluate the participants' performances on the basis of the standard measures such as trends in the reaction time.

C. Experimental Procedures Utilized in Implicit Learning Studies

As it was stated previously, Reber (1993) suggested two phases for grammar learning studies: the acquisition phase and the testing phase. Acquisition phase was the stage at which participants acquired some knowledge of the rules underlying the artificial grammar. In testing phase participants' level of understanding was measured by giving them new strings that were or were not based on the same rules of grammar. It is quite possible to generalize this staged view of a grammar learning study to other studies, maybe by adding an initial briefing phase. The following discussion describes this staged view of an implicit learning study.

1. Briefing Phase

A typical study of implicit learning will start with a "briefing phase" in which participants will be given instructions relating to the task that they will perform. As Smith (1999) noted, manipulation of the instructions given to participants will help observe the differences between the performances of the groups of participants receiving different sets of instructions. However, the unreliability of methods of measuring learning (such as post-task questionnaires) makes it difficult to search for possible correlations between the explicitness level of instructions and the explicitness level of participants' understanding of tasks. Instructions given during a briefing phase may only be useful to detect any positive or negative effects of explicit instructions on implicit learning. For example, Jimenez et al.

(1996) used two subject groups to show that probabilistic sequence learning is independent of participants' intention to learn. They did that by explicitly telling one group of participants about the relationship between consecutive locations of a target entity that they would periodically see on the computer screen.

2. Experimental Phase

This is an extension of what Reber (1993) called the "acquisition phase." At this stage, a subject is expected to acquire some implicit knowledge about the rules underlying the experimental task. The type of stimuli given to the participants depends on the type of the task employed in the study. In artificial grammar learning tasks, the stimuli will be sequences of letters presented as letter strings. In control tasks, the stimuli will be unexpected changes in the values of input or output parameters. In sequence learning tasks, the stimuli will be visual cues on computer screens, such as numbers or bright spots. Regardless of the type of the stimuli, the experimental phase of an implicit learning study will consist of repetitive trials. The number of trials will affect the statistical reliability of the results. The rules that the participants are expected to learn will be hidden in the grammatical structures of letter strings, relationships between the input and output parameters of simulations, and in the relationships between consecutive stimuli of probabilistic sequences.

It should be noted that experimental phase is not always a stage at which participants passively acquire an understanding of the system. This may be true for some artificial grammar studies, but experimental phase of a typical implicit learning study almost always involves recording of user responses. The changes in the responses help

detect the trends in users' reactions. In their sugar production control study, Berry and Broadbent (1984) recorded the users' actions and the output levels that they achieved. Lewicki et al. (1988) recorded the reaction times of participants. In their experiment, the reaction time was the time it took for a subject to detect the visual stimulus that was present in every trial.

3. Testing Phase

In grammar learning studies described by Reber (1993) this is the stage at which participants are asked to perform the actual task by utilizing the knowledge that they gained during the acquisition phase. However, testing phase does not always require a subject to apply acquired knowledge to previously unencountered stimuli. It may consist of giving post-task questionnaires to participants to determine the level of their awareness of the rules that they learned. As it was mentioned before, no reliable measuring method exists to determine the awareness level. Therefore, more reliable testing methods may be necessary.

One method is to look for performance hits caused by changes in the rules of the tasks. The use of different sets of letter strings by Reber (1969) is an example of that. He observed how the performance of an experimental subject changed from one set to another and determined that the performance of a typical subject was unaffected if the two sets of strings used the same grammar rules. In a sequence learning task, Lewicki et al. (1988) altered the sequence of stimuli after a certain trial and they observed sharply increased latencies in subject responses.

Chapter 2

IMPLICIT LEARNING IN PERCEPTION OF MUSIC

Studies done under the music perception mostly dealt with structure inherent in music. The main concern in these studies was to understand how knowledge of Western music regularities represented and learned through passive exposure to musical exemplars. This is because they did not aim at determining the role of implicit learning, though they often acknowledged some form of relationships were unknowingly established by listeners of music, without being exposed to some sort of formal training. Neither did they approach to perception of music as a type of pattern recognition, though they were seeking to understand the roles of structural elements like tone, pitch and rhythm in forming musical patterns differing in tonal hierarchy and harmony. In fact, the expectancy studies mentioned above point out to the fact that musical chords and melodies are perceived as complete patterns and any missing element invoke a sense of incompleteness. One may only note an exception in the sense that the completion of such an incomplete pattern can be achieved not in a predetermined way, but in a number of different ways, depending on the relationships of hierarchy and harmony.

In sum, studies almost always dealt with prior learned regularities and expectancies in auditory domain and left the concept learning of new information relatively unexplored. On the other hand, implicit learning experiments in visual area mostly dealt with learning of newly presented information. They found that people could also learn previously

unknown information as easily as they could learn previously established regularities, if the unknown information had a predictive role. An implicit learning study in auditory domain would also involve the same question that was asked in visual area, namely, whether or not people could rapidly learn predictive information through auditory patterns as they learned with visual patterns.

IV. A BACKGROUND ON THE STRUCTURE OF MUSIC

As Gabrielsson (1993) noted, music (specifically, Western music) is composed of many different elements, such as timbre, rhythm, melody, and harmony. Being the “tonal attribute of sound” (Lundin, 1967), timbre is the most essential element according to Gabrielsson. Other elements follow timbre in the order of importance, though they almost always contribute to the structure of music. It is not possible to individually explain all these elements in detail without deviating from the main stream of this paper. However, it will be useful to discuss the parallels between the structure of music and the structure of a typical implicit learning task.

A. Music as a Complex Structure of Sound and Tones

Pierce (1999) argues that the regularity of sound frequencies produced by strings of proportional lengths must have been noticed in ancient times by people of different cultures. It is quite likely that these independent discoveries led to the integer ratios that dominated the music scales developed by those cultures. In the last few centuries, mathematicians and music experts found that musical sounds as well as arbitrary sound

events could be replicated by combining different intensities of periodic frequencies. As Pierce explains, the frequency components (called “overtones”) produced by music instruments are all multiples of some fundamental frequency. The variations in relative intensities of the overtones define the “timbre” associated with the particular instrument producing the musical sound. Pierce describes the intensity variations of frequency components over time as “the general spectral shape” and he emphasizes that it is “crucial to timbral quality.” In summary, both frequency combinations and intensity variations carries information that helps identify the physical source of sound. This is especially true for music.

Of course, the above discussion relates to perception of sound in general, rather than perception of music in particular. However, the almost-universal acceptance of music scales implies that there is more to perception of music than the ability to distinguish frequency combinations. In his discussion on intervals and scales, Burns (1999) describes the standard Western music scale as a “12-tone chromatic scale of equal temperament.” He argues that other cultures’ scales are equivalent to Western scale in practice, even if they contain more than 12 intervals in theory. A detailed discussion on the comparison of music scales is beyond the scope of this paper, so it will suffice to say that music scales are based on the perceived equivalence of tones differing by a certain frequency interval called an “octave.” Music scales generally divide each octave into frequency intervals that have identical widths in a logarithmic scale.

There is a wealth of studies on how perception of music depends on the positions of tones in the music scale. Dowling (1999) summarizes certain papers that studied under

what conditions changes in familiar melodies made them unrecognizable to listeners. He describes the “patterns of ups and downs” of musical tones as the “contour” of a melody and lists evidences that deviations from the contour of a melody are easily caught by listeners. In other words, the contour is the most identifying aspect of music. According to Dowling, even infants recognize melodies on the basis of contour, but as one enters adulthood, one learns to recognize melodies on the basis of tonal structure. The perception of tonal structure is a significant advance; it enables listeners to distinguish melodies that are identical in contour but shifted in frequency. However, Dowling also quotes some evidence to show that small variations in individual tones (as much as a semitone, the smallest interval in the Western music scale) are not recognizable to many listeners. The secret is in the recognition of the “key” that the melody is based on. Musical melodies are not constructed by using all possible tones within an octave, but they are based on certain subsets of the tones, called “keys.” Dowling points out that while many listeners can recognize key changes, they are not able to detect variations in a melody that do not deviate from the key. Finally, Dowling presents results from other studies showing that even nonmusicians are influenced by tonality.

All this discussion indicates that listeners of music somehow perceive the rules employed by creators of music, even if they are not familiar with the structural aspects developed by musicians over ages. This may be explained as being the result of gradual diffusion of music knowledge across populations sharing the same heritage, but that would be a naïve conclusion. The theory of music has always been too complex even for many musicians to wholly learn and explain. We could conclude that there are perceptual and

learned aspects of music knowledge. Thus it is clear that some innate mechanisms is at work in perception of music, just like in implicit learning

B. Music as a Form of Grammar

The claim of this paper is that, the widespread use of discrete tones is an indication of a language-like structure embedded in music. Burns (1999) likens the practice of using a specific set of tones in music to the tendency of using a specific set of sound elements in auditory communication between animals and humans. He also points to religious and ritual sources of music as possible reasons for using a standard framework in music developed by any group of people. Moving from the analogy of music to language, researchers have often developed models to represent the music as a form of alphabet or as a grammatical structure (West et al., 1985). Although West et al. argue that “music does not exist to convey meaning” like the spoken (or written) language does, but they also admit that “much music is intended to express ideas or portray events.” They summarize Lerdahl and Jackendoff’s works (1977, 1982) as one of the most successful models of musical structure developed using the grammar analogy. That model dealt with four kinds of structural analysis to music and applied the well-formedness rules that are the basis of grammatical structures.

Carterette and Kendall (1999) agree that music has “syntax” just as a language does. When it comes to the explaining what the meaning of music is, they list several studies suggesting that emotions invoked by music may be viewed as the meaningful content of music. Countless studies that they quote indicate that emotions intended to be

conveyed by melodies sharing the same tonal structure is perceived in the same way by people of different cultures. In terms of its emotional content, music may be – somewhat naively- thought to be a universal language, but many critics have noted that emotional responses to music are “determined by cultural tradition rather than by the inherent qualities of music” (Gregory and Varney, 1996). The multitude of structural elements of music also makes it too ambiguous compared to a language. Aiello (1994) indicates that “the syntax of music has much more latitude than that of language.”

In summary, experts agree about the existence of a grammar-like structure in music. From the viewpoint of this paper, that should be enough to conclude that perception of music can be studied like the perception of grammar. However, as the above discussion implies, there is no consensus on if any real meaning is carried by music. The goal of the study presented here was to convey some hidden information (not some emotion) through certain elements embedded in melodies. There seems to be very few studies done to show how meaningful information can be encoded by employing elements of musical structure. The relationships between musical elements should lead to some expectations on what elements will follow those that are heard. Meyer (1994) argues that it is that expectation that embodies the meaning of music. In other words, grammar of music, if one is familiar with it, enables one to predict how a melody will proceed. This is an important point because it implies that the information carried by music is self-limited and must not violate self-consistency of music. This implication must be remembered when it is time to evaluate the findings of this study.

C. Music as a Collection of Patterns

Gabrielsson (1993) views the perception of rhythm as the perception of “grouping of the sound events,” or as a pattern. Sound patterns that make up the rhythm repeat themselves in a regular manner. Regularity of patterns is perceived through the periodic occurrences of “accented” (or, emphasized) members of patterns. As Butler (1992) noted, perceptual grouping of sound tones into patterns invoked a series of studies trying to apply Gestalt grouping principles to the perception of music. According to the results of some studies summarized by Butler, perception of grouping was not uniform from one melody to another, or even among the different performances of the same melody. As Van Noorden (1975) found, listeners grouped tones according to pitch proximity (i.e., according to how close the frequencies were), but as the tempo got slower even tones highly differing in pitch could be grouped together. Dowling (1973) found other interesting results by “interleaving” two melodies (by merging them into one melody). The listeners could easily recognize at least one of the melodies if the notes of the merged melodies sufficiently differed in pitch. There are, of course, many more studies with different results, but they cannot be discussed here in detail. Even this short discussion is sufficient to conclude that listeners tend to group elements of music.

Deutch (1999) argues that auditory system relies on grouping to reconstruct the original sound events from the mixture of auditory stimuli that it receives from different sources. As the above mentioned studies imply, the auditory system has the ability to reconstruct not only the original sound elements, but also the structure of musical melodies, even if they are intertwined with other melodies. However, as Deutch notes, all those

different characteristics of music cannot be perceived by employing a single grouping mechanism. She suggests that more than one grouping mechanism at work and they employ different dimensions such as frequency, amplitude, timbre, etc. However, some of those dimensions may be more dominant than others. Jones (1993) presents evidence that average listeners compares the “dynamic shapes” while trying to detect a similarity between two melodies. However, she emphasizes that dynamic shape is not simply pitch contour, i.e., it is not only dependent on ups and downs of frequencies of tones. The timing of the accented tones and the rhythm complements the contour in defining the dynamic shape. This last point is an important one, because it implies that shifts in individual tones of a melody may escape detection if the overall shape of the melody is preserved. This implication will also be important in evaluating the results of the study presented here.

V. IMPLICIT LEARNING IN PERCEPTION OF MUSIC

Davies (1978) argues that music “does not really satisfy the requirements that would justify its being called a language,” but he also notes that common people have little difficulty in distinguishing music from other collections of sound events. The fact that everyday sounds are labeled as mere “noise” and judged to be different from music implies that music has a content other than sound waves that it is composed of. However, again as Davies notes, meaning of music cannot be explained, even though it conveys some emotion to any sentient listener. The ability to derive emotions from music is in strict contrast with the inability to explain the meaningfulness of music. This contrast reminds one of the

absence of verbalizable knowledge derived from an implicit learning task. The implication is that implicit learning may play a role in perception of music. This part will briefly discuss the evidences for that role.

A. Pattern Recognition in Music Perception Studies

As it was mentioned earlier, a strong evidence for pattern recognition in perception of music is in the recognition of a melody preserving its contour even if it is played in different keys. According to Bartlett (1993), two familiar American melodies were recognizable to listeners even if they were played in 12 different keys. It was also mentioned that contour was not the only element employed by listeners looking for patterns in musical melodies. Rhythm, pitch proximity were also counted among important pattern-forming factors. Complexity of interrelationships between those musical elements should make musical patterns ideal tools for implicit learning experiments emphasizing pattern recognition. However, all of the studies mentioned above studied pattern recognition only in the context of recognizability of musical melodies. They did not approach the problem from the perspective of implicit learning. In fact, only a few studies investigated the implicit learning directly with strings of musical events, and then the strings were not real melodies but only arbitrary tone sequences (Bigand, et al., 1998). It is clear that much needs to be done in implicit learning in perception of musical patterns.

Tonal music employs patterns of more or less equal length called "meter." Therefore, most melodies can be considered to be collections of patterns. Thus it is quite possible that listeners perceive the complex principles of music through the regularities

perceived in musical patterns. The fact is, no listener –except learners of music– is given an explicit description of musical principles. Any understanding must be gained implicitly, and that may be occurring through establishing commonalities between patterns forming a complete melody. Evidence exists that nonmusicians with plenty of listening experience can even identify composers of melodies that they did not hear before. They clearly do that by looking for the patterns that they previously heard in their favorite composers' known works. The viewpoint of the present study is that, musical patterns not only help listeners implicitly understand the underlying principles of music, but they may also help listeners perceive non-musical information hidden in melodies.

B. Expectancy in Perception of Music Perception Studies

Implicit or explicit knowledge of music of listeners will influence their expectations on what musical element will follow those already heard. In his discussion on mental representations of musical relationships, Butler (1992) presents evidence that listeners develop expectations on the rhythms or tone sequences after hearing several musical patterns sharing certain characteristics. Butler also notes that listener expected a well-known tune be played with the tempo with which they are most familiar, even if they are hearing the melody for the first time in a study. In a review of studies employing rhythm and meter, Butler indicates that irrespective of their formal training in music, listeners could reproduce the rhythms of even unfamiliar melodies after hearing portions of them, provided that the melodies employed common rhythms in Western music.

In summary, listeners of music are able to store information on fundamental characteristics of pieces of melodies and utilize that information to correctly estimate the characteristics of complete melodies. It is clear that information on frequently heard melodies are stored in long-term memory, because they lead to expectations on future playbacks of those melodies even if the original versions have not been heard for a while. The above mentioned ability to replicate commonly occurring rhythms implies that knowledge of certain musical elements like rhythms is stored independently of the information on specific melodies. Thus it can be concluded that participants implicitly learn about the structure of a previously heard melody or about the structure of music in general and they retain that knowledge in memory. The studies mentioned earlier studied expectancy in musical perception to determine the listener's ability to complete the tone or rhythm patterns of actual randomly constructed melodies. However, it should also be possible to prompt participants of experiments to develop expectancies of certain characteristics of musical melodies through auditory clues embedded in melodies. The predictive context of the present study is intended to do just that. By employing different types of clues made up of specific tone sequences or rhythm patterns, the present study is expected to answer questions on which music elements tell the participants what to expect, and to what degree.

C. Lack of Studies on Implicit Learning in Music Perception

The demonstrations of pattern recognition and the role of expectancy in studies mentioned above all point to the possibility that implicit learning plays an important role in

music perception. However, none of the researchers who did those studies referred to implicit learning in their conclusions. Among the authors cited so far, only Dowling (1993) specifically mentions the growing interest of psychologists on “automatic” learning that occurs without conscious intent and discusses the possible applications of this type of learning to models of music cognition. His terminology is different, but he clearly refers to implicit learning. Cuddy (1993) mentions the ability of listeners to abstract relationships between the tones without conscious effort. This is another indirect reference to implicit learning.

In summary, the lack of references to implicit learning in music perception should not be taken to mean that music researchers dismiss the concept of implicit learning. Their indirect references provide sufficient proof that they acknowledge the role of implicit learning in music perception. They only seem to be reluctant to make use of one of the most commonly used concepts in studies of perception of cognition. The growing interest toward implicit learning will no doubt also influence the researchers of music perception. However, apparent commonalities between music perception studies and implicit learning studies indicate that researchers from both sides have taken too long to show interest in each others’ domains. The present study is an effort to place music melodies in the context of an implicit learning task and to use common themes of implicit learning tasks in the context of a music perception study.

Chapter 3

A REVIEW OF THE PRESENT STUDY

VI. A GENERAL DESCRIPTION OF THE EXPERIMENTS

A. The Purpose of the Experiments

This paper has so far analyzed the role of implicit learning in the perception of musical structure. As it was discussed in the previous chapter, most of the participants of music perception studies demonstrated an unconscious awareness of musical structure in actual melodies, although many of them had no prior knowledge or training of music. This is the reason why almost all the studies on the perception of music structure have come to assume that the performance of a typical participant was influenced by his implicit knowledge of music. In other words, implicit learning has been accepted as the primary mechanism of learning of music for the listeners who lack a formal training and maybe even for those who have had formal training. The prominence of implicit learning in music perception suggests that musical stimuli may also help study implicit learning in situations other than passive listening of music. However, the literature of implicit learning shows a lack of studies employing music as the primary carrier of information. Although the subject of implicit learning has been extensively studied with finite strings of symbols and visual patterns, only a few studies investigated the implicit learning directly with strings of musical events. The structure of music is much more complex than the finite state

grammars used in implicit learning studies, and this structural complexity makes musical patterns suitable to be used as stimuli in implicit learning experiments. Since the formal knowledge of music is available only to a select few, actual musical melodies are complex enough for most people. Thus, experimenters do not have to construct artificial musical patterns to be employed in implicit learning studies.

Therefore, the study presented in this paper aimed at examining the implicit learning in auditory perception by using auditory stimuli, namely, musical melodies. The purpose was to determine how certain structural elements of music could be utilized to invoke implicit learning of probabilistic relationships between those elements and the information needed to perform an experimental task. It was assumed that the complexity of structure in musical stimuli would ensure that the performance of the experimental task would be improved only through implicit learning of the probabilistic relationships. The task performance was measured by the usual criteria, such as the rate of success and reaction time. The study consisted of four experiments in which different combinations of structural elements were employed as contextual cues to promote implicit learning by participants.

B. General Method of the Experiments

1. A Description of the Experimental Task

In every experiment, the task was to decide whether or not two melodies played back to back were the same. The two melodies formed a “same-pair” if they were identical to each other. If one note of the second melody differed from the corresponding note of the

otherwise-identical first melody, the two melodies formed a “different-pair.” The difference between those notes was a shift of a semitone or a full tone. In other words, the difference between the two melodies of a pair was a difference in the pitch of a single note.

2. The Method of Administration

In all four experiments, a complete experimental session consisted of 150 trials preceded by a training period lasting 10 trials. In each trial, two melodies forming a pair were played back to back by a computer application. The second melody was played back exactly one second after the first melody and waited for a response from the listening participant during and after the playback. The participant entered a response by pressing one of the two keys on the keyboard: ‘S’ for a “Same” response, ‘L’ for a “Different” response. In order to force the participant to listen to the second melody of the pair before entering a response, the application disabled the keyboard of the computer until the second melody started playing. The keyboard was re-enabled once the second melody of a pair started; thus the participant was able to respond as soon as she made a same/different decision. The time that passed from the beginning of the second melody until the participant made an entry was recorded as the reaction time of the participant.

An important feature of all these experiment was the feedback given after each trial. A verbal message indicated whether the response of a participant given to each trial was “Correct” or “Wrong.” These messages were played back from digital audio files by the computer application just like the melody stimuli. The feedback messages were intended to give directions to participants and thus to promote implicit learning. This was not done in most implicit learning experiments based on relatively straightforward tasks in which

participants could discover their success through nonverbal, indirect clues.

However, in the auditory perception tasks discussed in this paper, there were neither any visual clues, nor any indirect means that would signal the success of a particular action to a participant. Feedback messages were therefore necessary to help participants discover the appropriateness of their responses. The feedback also added a limited interactivity to the experimental task and thus kept the attention of the participants active.

3. Experimental Stimuli

Melodies utilized in the experiments were short excerpts from various international folk music or classical music melodies. Melodies were obscure enough to reduce the chances that a participant could succeed in the task through his familiarity with the melodies. Melodies were converted into digital format by a music transcription software. Resulting digital melodies were all of equal length. All four experiments used the same set of 150 melodies. In each trial, the computer application randomly decided whether or not a pair of melodies would be a same-pair or a different-pair. For a same-pair, the application played a melody twice. For a different-pair, the second melody was a modified copy of the first one in which a single note was shifted by a full tone or by a semitone. The numbers of same-pairs and different-pairs were the same. In each session there were 75 same-pairs and 75 different-pairs.

4. The Follow-up Questionnaire

Following the completion of all 150 trials of an experimental session, a participant was asked to complete a questionnaire to help assess several aspects of her awareness of the experimental task and the rules of the clues embedded in musical stimuli. A copy of the

complete questionnaire is presented in Appendix F (page 165). The primary purpose of the questionnaire was to determine whether or not a participant's learning was implicit. Therefore, the first question of the questionnaire inquired if the participant's deliberately looked for a relationship or a rule that would facilitate the performance of the task. If the answer to that question was affirmative, the participant was asked to describe the relationship that she found and how she tried to utilize that relationship. If a participant closely described the probabilistic relationships represented by the predictiveness relationships, she would be deemed to have acquired the knowledge explicitly. The second question inquired if the participant noticed a characteristic of each melody pair that would indicate whether the two patterns were same or different. Again, an affirmative answer required the participant to describe the characteristic that she discovered and the methods of utilizing that information. The participant was then asked in which part of the experiment, i.e., whether at the beginning, in the middle or toward the end, she was more successful. Final questions asked the participant to evaluate the experimental task. These last questions were intended to obtain clues on how the future experiments of the same kind should be designed to ensure willing performance of participants unhindered by negative aspects that made the task especially tedious.

C. The Predictive Context of the Experiments

1. The Purpose of Clues

The beginning part of each melody served as a clue that predicted whether or not the melody pair was more likely to be a same-pair or a different-pair. Different clues had

different predictiveness relationships, meaning that each one represented a certain probability that a melody pair would be a same-pair or a different pair. It was this relationship that the participants were expected to learn implicitly. If a participant acquired the knowledge of predictiveness relationships of the clues, he could start responding faster and more accurately.

As it was implied in the background discussion on the structure of music, two structural elements of music contribute the most to perception of music: tone and rhythm. Therefore, two types of clues were used in experiments: melody clues and rhythm clues. Melody clues were sequences of unchanging notes with unchanging lengths. Rhythm clues were sequences of a certain number of notes with unchanging lengths. The rates of implicit learning achieved with those two types of clues would show the relative effectiveness of those two types of clues.

2. Predictiveness Relationships of Clues to Same-Trials

In each experiment, three distinct clues were used with different levels of association with same-trials. Positive clue predicted a same-trial with a probability of 75%. Thus positive clue was strongly associated with same-trials. Negative clue predicted a same-trial with a 25% probability; this clue was associated with different-trials. For neutral clue, the probability of a same-trial was 50%, meaning that this clue was not associated with either same-trials or different-trials. The relationship between positive clue and same-trials, or the relationship between negative clue and different-trials could be implicitly learned and exploited by participants to enter more and more correct responses with gradually decreasing reaction times.

3. Melody Clues

In the first two experiments there were three distinct note sequences that served as clues. Those were called “melody clues” because each clue was a short melody by itself. Each melody clue contained four or five notes, and they all had the same duration.

The following note sequence (see Figure 1) served as the positive melody clue. As it was explained above, 75% of melody pairs that started with positive clues were same-pairs.



Figure 1 The positive melody clue used in Experiments 1 and 2

The following note sequence (see Figure 2) served as the negative melody clue. Only 25% of melody pairs that started with negative clues were same-pairs.



Figure 2 The negative melody clue used in Experiments 1 and 2

The following note sequence (see Figure 3) was used as the neutral melody clue. Half of the melody pairs that started with neutral clues were same-pairs and the other half were different-pairs.



Figure 3 The neutral melody clue used in Experiments 1 and 2

The following figure shows the notes of a melody pair starting with the *neutral* melody clue. The note which is shifted (by a semitone) in the second melody of the pair is enclosed in a box.



Figure 4 A melody pair that starts with the neutral melody clue

4. Rhythm Clues

In the third and fourth experiments, the first three notes of each melody followed one of the three rhythm patterns that were called the “rhythm clues.” The actual notes in a rhythm clue could be different in any of the melodies with the same clue. However, the relative lengths of the notes in a rhythm clue did not change.

The positive rhythm clue consisted of a fourth note followed by two eighth notes:



Figure 5 The positive rhythm clue used in Experiments 3 and 4

The negative rhythm clue had two eighth notes followed by a fourth note:



Figure 6 The negative rhythm clue used in Experiments 3 and 4

Finally, the neutral rhythm clue consisted of three fourth notes:



Figure 7 The neutral rhythm clue used in Experiments 3 and 4

D. Reversal and Nonreversal groups

1. Reversal of the Predictiveness Relationships

In order to determine whether or not a participant responded to trials according to the predictiveness relationships of clues, those relationships were changed in some experimental sessions. It was expected that the performance of a participant would be adversely affected by that change, if the participant tended to respond according to the predictiveness relationships of the clues after learning them implicitly.

The change consisted of reversing the predictiveness relationships of positive and negative clues after the 90th trial of a session. Therefore, such a session was called a “reversal session.” After the 90th trial of a reversal session, the positive clue made it 75% likely that a pair would be a different-pair, i.e., it served like the negative clue did before

the 90th trial. The negative clue, on the other hand, made a pair 75% likely to be a same-pair. The neutral clue's predictiveness level did not change in a reversal session.

2. Differences between Reversal and Nonreversal Groups

The participants who undertook reversal sessions formed the reversal groups in each of the four experiments. The performance of a reversal-group participant was expected to undergo a dramatic change after the reversal of the predictiveness relationships. However, this conclusion could not be justified without eliminating the effects of other factors that may have contributed to performance changes. Thus a nonreversal group was necessary to determine the changes in an average participant's performance in the absence of a reversal. Therefore, some participants undertook "nonreversal sessions" where the predictiveness relationships did not change for any of the clues. Those participants formed the nonreversal groups of the four experiments. The performance of an average nonreversal-group participant was used as a reference point while analyzing the performances of reversal-group participants.

The following tables summarize the makeup of reversal and non-reversal sessions undertaken by reversal-group and nonreversal-group participants:

Table 1 Numbers of Same/Different Pairs in a Reversal Session

<u>Trial Block</u>	<u>Positive Clue</u>	<u>Negative Clue</u>	<u>Neutral Clue</u>
1-90	30 pairs	30 pairs	30 pairs
	23 same	7 same	15 same
	7 different	23 different	15 different
91-150	20 pairs	20 pairs	20 pairs
	5 same	15 same/	10 same
	15 different	5 different	10 different

Table 2 Numbers of Same/Different Pairs in a Non-Reversal Session

<u>Trial Block</u>	<u>Positive Clue</u>	<u>Negative Clue</u>	<u>Neutral Clue</u>
1-90	30 pairs	30 pairs	30 pairs
	23 same	7 same	15 same
	7 different	23 different	15 different
91-150	20 pairs	20 pairs	20 pairs
	15 same	5 same/	10 same
	5 different	15 different	10 different

E. An Overview of Experimental Data

1. Independent Variables of the Experiments

a) Type of the Clue

The purpose of the experiments was stated as examining implicit learning through auditory stimuli. As it was mentioned before, different structural elements of music were expected invoke implicit learning in different degrees. Thus the primary independent variable was the type of the clue, namely, whether it was a melody clue or a rhythm clue. Experiments 1 and 2 employed melody clues and 3 and 4 employed rhythm clues. The knowledge that would be acquired by implicit learning was the probability relationships between the clues and the melody pairs. Therefore, another important variable was the predictiveness relationship of a clue with same-trials, i.e., whether it signaled a same-trial with a 75%, a 50%, or a 25% probability.

b) Playback Tempo

As it was noted before, melodies used as stimuli were selected from a wide range of obscure musical works of highly different styles. They represented a highly diverse sample of various combinations of structural elements, such as key and contour. Their playback tempo, however, could be controlled to determine the influence of that element in implicit learning. It was predicted that a typical participant could distinguish the individual notes of a melody played at a slow tempo. In contrast, a faster tempo would probably cause a melody to be perceived as a complete pattern in which individual notes would not stand out. Thus the playback tempo was the secondary independent variable. For either one of the

two types of clues, melody stimuli were played back at a slow tempo (120 beats per minute, meaning that the duration of a standard note would be $1/120$ of a minute) in one experiment, but at a faster tempo (200 beats per minute, i.e., $1/200$ minute per standard note) in the other experiments. Thus a different combination of two types of clues and two tempo choices were employed in each of the four experiments.

2. Performance Measures (Dependent Variables) Targeted by Analysis

a) Accuracy Score

Data analysis focuses on two potential performance measures for a participant: the overall accuracy of the participant in classifying the pairs as same or different, and his average reaction time in giving his responses. The former measure, i.e. the ratio of the correct responses, was called the "accuracy score." The accuracy score was also a criterion used to eliminate the participants who made no effort to perform the given task or who could not perceive the structure of the stimuli. Accuracy score of a randomly-performing participant would be around 50%. It can be said that the accuracy scores would form a binomial distribution if all participants responded randomly all the time. Using the binomial distribution formulas, it is easy to show that the probability of obtaining a score above 54% by chance is less than 20%. Thus a limit of 54% was set to exclude the great majority of participants who did not perform beyond the chance level.

b) Reaction Time

The second important performance measure was reaction time. For the purposes of the current study, reaction time was the time interval between the beginning of the second

melody in a pair and the moment the participant entered a response. In that time period, the participant was assumed to be mentally processing the auditory stimuli, starting with the auditory clue embedded in the melody.

F. An Overview of Data Analysis Methods

1. Effects of Implicit Learning on Performance

As it was mentioned above, the average accuracy score was the most important performance measure. However, an average score for the whole session undertaken by one participant could only tell the participant's overall ability to detect the presence or the absence of subtle tonal differences between pairs of melodies. In order to determine if a participant acquired the knowledge of the task implicitly, it was necessary to calculate and compare the average scores for separate groups of trials between which there should be differences arising from the influences of implicit learning.

Implicit learning was expected to show its influence on the task performance in three ways: 1) An improvement of performance during the first 90 trials and 2) Pattern of performance as a function of the clue-trial type combination. 3) A performance cost resulting from the reversal of the predictiveness relationships of clues in a reversal session.

In order to determine if implicit learning improved the task performance, the mean values of the performance measures (i.e., the accuracy score and the reaction time) were analyzed before the 90th trial for both experiment and nonreversal groups. Primarily, we expected a better performance on trials with congruent relationships between the clue type and trial type. For instance, performance on same-trials having positive clue should be

better than performance on same-trials with negative clue. The performance cost resulting from the reversal of the predictiveness relationships was analyzed by comparing the mean performance measures before and after the 90th trial. That analysis was done separately for experiment and nonreversal groups, because significant changes were expected for reversal-group participants only. The nonreversal-group participants were expected to maintain or further improve their levels of performance. All those analyses required tracking the changes in the mean values of performance measures from one trial block to another, as explained below.

2. Analysis of Performance Changes before the 90th Trial

In order to detect the changes in the task performance before the 90th trial, the trials until the 90th were split into two equal blocks of 45 trials each. The accuracy scores and reaction times of all participants of the same experiment were averaged separately for these 45-trial blocks. The differences between the mean scores or mean reaction time of the two blocks showed the changes of an average participant of the experiment. Furthermore, since performance changes were expected to reflect the clue-trial type relationship, mean scores and reaction times were calculated separately for same and different-trials starting with positive, negative and neutral clues. As it is explained below, appropriate statistical tests were applied to determine whether the trial type (i.e., whether a pair was same or different), or the clue type (positive, negative or neutral) had significant effects on the variations of the experimental constructs. The same tests also helped determine whether the mean accuracy score or mean reaction time changed significantly from the first trial block (trials 1 to 45) to the second trial block (trials 46 to 90).

3. Analysis of Performance Changes after the 90th Trial

The reversal of the predictiveness relationships of clues was expected to lower the accuracy score of a reversal-group participant. Until the 90th trial, a reversal-group participant was expected to learn the task structure implicitly. The trials immediately after the 90th trial would contradict that learned structure and thus the participant could not succeed with what she had learned until the 90th trial. However, it was also possible that the participant could re-learn the changed task structure and improve her performance again. Therefore, in order to observe the performance change resulting only from the reversal of the predictiveness relationships, it was necessary to compare the mean accuracy scores and mean reaction times for trial blocks immediately before and after the 90th trial. The choice was to split the trials after the 90th into two equal blocks just like it was done for trials before the 90th trial. Thus the mean values of both experimental construct were calculated over four trial blocks:

- Block 1: Trials 1-45
- Block 2: Trials 46-90
- Block 3: Trials 91-120
- Block 4: Trials 121-150.

Performance changes resulting from the reversal of the predictiveness relationships were expected to reveal themselves through the differences between the mean values of 2nd and 3rd blocks. Of course, if the predictiveness relationships did not change for any clue, which was the case for nonreversal-group participants, then no significant changes would be observed between those two blocks. Therefore, performance changes between the

Blocks 2 and 3 were expected to be significant for reversal-group participants, but not for nonreversal-group participants.

In addition to comparing mean scores and reaction times from Blocks 2 and 3, mean values were also compared for Blocks 3 and 4, so that the effects of any re-learning could be detected. In other words, changes in the task performance after the 90th trial were examined under two headings: 1) Changes due to the reversal of the predictiveness relationships. 2) Changes due to the re-learning of the predictiveness relationships.

4. Expected Performance Changes With Respect to Clues

In summary, the average values performance measures (i.e., accuracy score and reaction time) were calculated separately for each of the four trial blocks. Within each block, means were calculated separately for same-trials starting with positive, negative and neutral clues, and for different-trials starting with positive, negative and neutral clues.

As it was noted before, an average participant's performance was expected to improve until the 90th trial for positive and negative clues because they helped one predict whether a melody pair would be the same or different. For example, because the positive clue predicted a same-trial, the average accuracy score for positive-same trials would be greater in Block 2 than in Block 1. Conversely, because the negative clue predicted a different-trial, the average accuracy score for negative-different trials would also increase from the first block to the second block. On the other hand, the accuracy score for positive-different trials would probably drop from Block 1 to Block 2, because they would contradict the expectations a participant would have about the positive clue. Similarly, the accuracy score would also drop for negative-same trials contradicting the predictive context

of the negative clue. The accuracy score would be likely to remain unchanged for any trials starting with the neutral clue, since that clue equally predicted a same or a different pair.

The performance cost of the predictive context reversal was expected to cause a drop in the average accuracy score of a participant, but a relearning of the new reversed relationships was also possible. And that would reduce the negative effects of the reversal. Score drop, if observed, would be more apparent between the trial blocks immediately before and after the 90th trial, i.e., between trial blocks 2 and 3. However, it was possible that a rather fast relearning of changed relationships would reduce the score drop or prevent it. It was also possible that the reversal of the predictiveness relationships would only result in un-learning of older relationships; changes in the task performance would be completely unpredictable in such cases.

Average reaction times were also expected to be affected by the learning or the reversal of predictiveness relationships, but very little variations were observed in reaction time with respect to trial blocks or different clues; therefore, only the mean accuracy scores were analyzed to detect the presence of implicit learning.

5. Statistical Tests Utilized in Analyses

The effects of trial block and clue type on accuracy score and reaction time were analyzed using repeated-measures ANOVA test. Where applicable, main effects and interactions of trial block and clue type on accuracy score are presented in the following sections. However, because reaction time did not show sensitivity to either the trial block or

the clue type, only the mean values are presented for that variable without the results of statistical tests (see Appendices B, C, D, and E).

Chapter 4

EXPERIMENT 1

VII. OVERVIEW OF EXPERIMENT 1

A. Purpose

The purpose of Experiment 1 was to examine the process of implicit learning in auditory perception by utilizing the predictive context of melody clues embedded in musical melodies.

B. Method

The main task in Experiment 1 was the same task commonly used throughout the present study: determining whether or not a pair of melodies was same or different. Melody clues, i.e., specific tone sequences embedded in melodies provided predictive information relating to whether or not a specific pair was likely to be a same-pair or a different-pair. Three different melody clues (see page 34) of three different predictive levels (75%, 50% and 25%) were used in equal proportions throughout 150 trials that made up each experimental session. In Experiment 1, melodies were played at a slow tempo (120 beats per minute; one standard note lasts 1/120 of a minute) to enable the participants to better distinguish the individual notes. Each melody (including its clue) lasted about 7 seconds.

C. Participants

A total of 71 people participated in this experiment, 44 in the reversal group, and 27 in the nonreversal group. Participants with overall accuracy lower than 54% (see the discussion in page 40) were considered to have responded mostly randomly and their data were not included in the analysis. The data from 35 reversal-group participants and 16 nonreversal-group participants were included in the analysis because their overall accuracy scores until the reversal of the roles of clues were above the random-performance limit.

VIII. RESULTS OF EXPERIMENT 1

A. Analysis of Performance Changes before the 90th Trial

1. Mean Scores of Same-Trials

The following table shows the summary statistics for the accuracy scores of Experiment 1 for same-trials:

Table 3 Means, Standard Deviations and N for Accuracy Scores for Same-Trials before 90

Same-Trials				
Clue	Trial Block	Accuracy Score (Mean)	Accuracy Score (Std. Dev.)	N
Positive	1-45	0.6867	0.1806	51
	46-90	0.6272	0.1749	51
Negative	1-45	0.5369	0.3192	51
	46-90	0.6301	0.2883	51
Neutral	1-45	0.6248	0.1519	51
	46-90	0.6642	0.1416	51

The following figure shows the changes in the mean accuracy scores summarized in the above table:

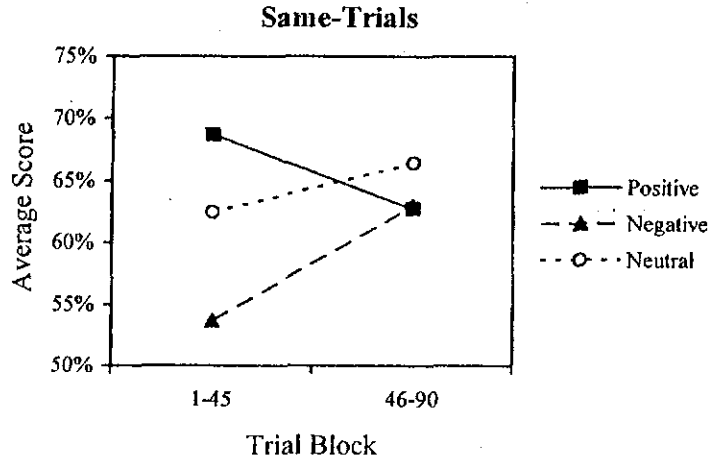


Figure 8 Mean Accuracy Scores of Same-Trials before 90

According to 3 (clue type) by 2 (trial block) repeated-measures ANOVA test, the expected performance improvements were not observed in same-trials. In fact, the mean accuracy score for positive clue decreased instead of increasing (mean = 0.6867 for trial block 1-45 and mean = 0.6272 for trial block 46-90). In addition, the average score for negative clue increased instead of decreasing (mean = 0.5369 for trial block 1-45 and mean = 0.6301 for trial block 46-90). There was a significant main effect of clue type ($F(2,100) = 4.194$, $MSe=3.759$, $p = 0.018$, $r = 0.27$). No significant main effect was found for trial block ($F(1,50) = 0.959$, $MSe = 4.743$, $p = 0.332$, $r = 0.13$). A significant interaction was found between the clue type and the trial block ($F(2,100) = 3.406$, $MSe = 4.482$, $p=0.037$, $r = 0.25$), although not in the expected direction. When neutral clue trials were eliminated from the analyses, the interaction was still significant ($F(1,50) = 5.453$, $MSe= 5.441$, $p =$

0.024, $r = 0.31$). Further analysis revealed the pattern of this interaction, that two types of trial blocks (1-45 and 46-90) worked equally well for Positive clue ($F(1,51)=2.177$, $p = 0.146$, $r = 0.20$) and for negative clue ($F(1,50)=2.290$, $p = 0.137$, $r = 0.21$). There was no significant main effect for trial block ($F(1,50) = 0.212$, $MSe = 6.856$, $p = 0.648$, $r = 0.06$). However, this lack of a main effect is the result of offsetting simple effects of trial block for the positive clue which showed a decrease (mean = 0.6817 to mean = 0.6317) in average score from block 1 to block 2 and of trial block for the negative clue which showed an increase in average score (mean = 0.5369 to mean = 0.6301) from block 1 and block 2. There was a significant main effect for clue type ($F(1,50) = 5.605$, $MSe = 4.909$, $p = 0.024$, $r = 0.31$). Overall, as hypothesized, accuracy score was higher for the positive clue (marginal mean = 0.657) than for the negative clue (marginal mean = 0.583). However, this effect is misleading because inspection of cell means and set of simple effects revealed that the average score for negative clue increased instead of decreasing and the mean accuracy score for positive clue decreased instead of increasing. Out of 51 participants, 23 participants' accuracy scores increased for positive clue. As for the negative clue, 28 participants' accuracy scores increased.

2. Mean Scores of Different-Trials

The summary statistics for the accuracy scores for different-trials are shown in the following table:

Table 4 Means, Standard Deviations and N for Accuracy Scores for Different-Trials before 90

Different-Trials				
Clue	Trial Block	Accuracy Score (Mean)	Accuracy Score (Std. Dev.)	N
Positive	1-45	0.5867	0.2666	51
	46-90	0.5917	0.2287	51
Negative	1-45	0.5923	0.1845	51
	46-90	0.6052	0.1537	51
Neutral	1-45	0.5314	0.3052	51
	46-90	0.6278	0.2776	51

The following figure shows the changes in the mean scores listed in the above table:

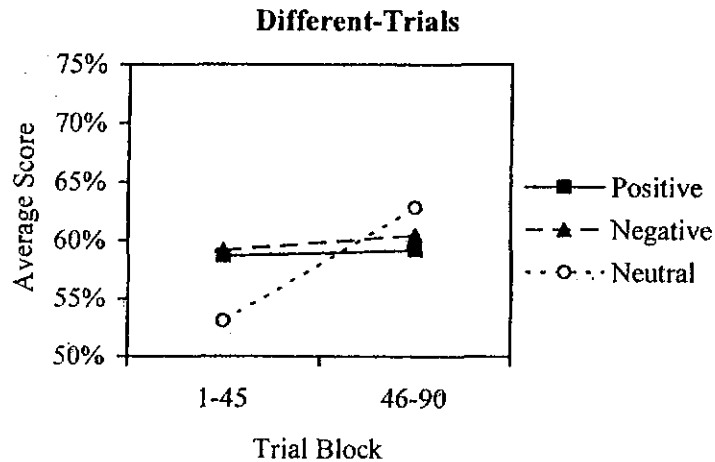


Figure 9 Mean Accuracy Scores of Different-Trials before 90

According to 3 (clue type) by 2 (trial block) repeated-measures ANOVA test, the expected performance improvements were not observed for different-trials. Mean score for positive clue was 0.5867 for the first trial block (1-45) and 0.5917 for the second trial block (46-90). In other words, they were not much different from each other, and there was a small increase whereas a decrease was expected. Mean score for negative clue was 0.5923 for the first trial block and 0.6052 for the second trial block, showing an increase as expected, but an insignificant one. ANOVA also confirmed that there was not a significant main effect of clue type ($F(2,100) = 0.213$, $MSe = 4.407$, $p = 0.809$, $r = 0.06$). Neither was there a significant main effect of trial block ($F(1,50) = 2.694$, $MSe = 4.130$, $p = 0.107$, $r = 0.22$). Finally, there was no significant interaction to be found between clue type and trial block ($F(2,100) = 1.359$, $MSe = 4.809$, $p = .262$, $r = 0.16$). When neutral clue trials were

eliminated from the analyses, the interaction was still not significant ($F(1,50) = 0.022$, $MSe = 3.610$, $p = 0.883$, $r = 0.02$). There was no significant main effect for clue type ($F(1,50) = 0.129$, $MSe = 3.609$, $p = 0.721$, $r = 0.05$). Overall, accuracy score was little lower for the positive clue (marginal mean = 0.589) than for the negative clue (marginal mean = 0.599). There was no significant main effect for trial block ($F(1,50) = 0.163$, $MSe = 2.536$, $p = 0.688$, $r = 0.05$). Overall, accuracy score was little lower for the block 1-45 (marginal mean = 0.589) than for the block 45-90 (marginal mean = 0.598). Out of 51 participants, 24 participants' accuracy scores increased for negative clue. As for the positive clue, 27 participants' accuracy scores increased.

3. Mean Scores of Positive-Clue Trials

In order to better focus on the changes occurring in the scores for positive clue, changes in the mean accuracy scores of same-trials and different-trials of positive clue were compared. As the following figure shows, the changes were in directions opposite to those expected. The mean score for positive-same trials was expected to increase from trial block 1 to 2, but it decreased. Very little change was observed in the mean score for positive-different trials, but it was expected to decrease.

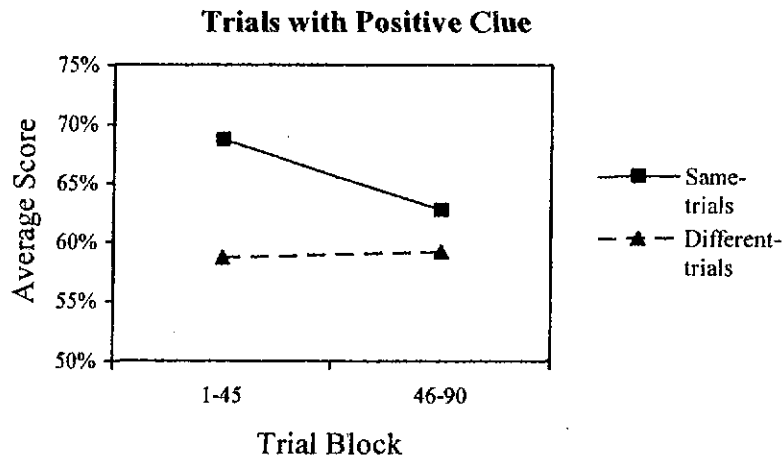


Figure 10 Mean Accuracy Scores of Positive Clue for Trials before 90

According to the repeated-measures ANOVA test, however, those changes, albeit in the wrong direction, were not significant. No significant interaction was found between trial type and trial block ($F(1,51) = 1.300$, $MSe = 2.837$, $p = 0.260$, $r = 0.15$). The test showed no significant main effect of trial block for trials starting with the positive clue ($F(1,51) = 0.636$, $MSe = 4.082$, $p = 0.429$, $r = 0.10$). Overall, accuracy score was higher for

trial block 1-45 (marginal mean= 0.635) than trial block 46-90 (marginal mean= 0.613). Also no significant main effect of trial type was found (i.e., same or different) ($F(1,51) = 2.629$, $MSe = 8.158$, $p = 0.111$, $r = 0.22$). Overall, accuracy score was higher for same trials (marginal mean= 0.656) with positive clue than for the different trials (marginal mean= 0.592).

4. Mean Scores of Negative-Clue Trials

A more detailed look at the changes in the mean accuracy scores of same-trials and different-trials of negative clue did not give the expected picture either. The following figure shows the changes in the mean score for negative-same and negative-different trials.

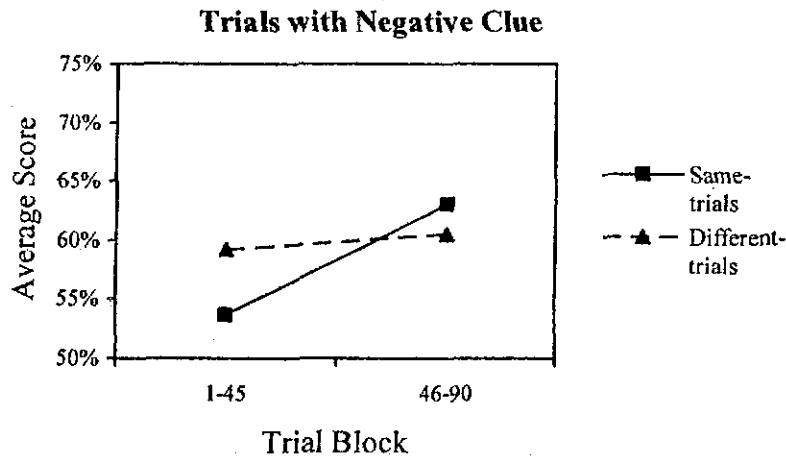


Figure 11 Mean Accuracy Scores of Negative Clue for Trials before 90

According to the repeated-measures ANOVA test the interaction between trial type and trial block was not significant ($F(1,51) = 3.171$, $MSe = 4.964$, $p = 0.081$, $r = 0.24$). ANOVA showed a significant main effect of trial block for trials starting with the negative

clue ($F(1,51) = 4.081$, $MSe = 5.747$, $p = 0.049$, $r = 0.27$). Accuracy score was lower (marginal mean = 0.560) for trial block 1-45 than for trial block 46-90 (marginal mean = 0.629). No significant main effect was found for trial type (i.e., same or different) either ($F(1,51) = 0.005$, $MSe = 7.792$, $p = 0.943$, $r = 0.00$). Overall, accuracy score was lower for same trials (marginal mean = 0.593) with negative clue than for different trials (marginal mean = 0.596).

5. Analysis of Performance Changes after the 90th Trial

The analysis of the performance changes before the 90th trial did not suggest any learning of probabilistic relationships between clue type and trial type. Therefore, the analysis of the task performance after the 90th trial was not found necessary to be included here. Tables, figures and discussions relating to trials after 90 are placed in Appendix A (page 140).

B. Summary of the Results

Judging from their mean accuracy scores, neither the reversal group participants nor the nonreversal group participants showed the expected performance patterns that would result from the implicit learning of the roles of clues.

The interaction between clue type and trial block was found to be significant for same trials, but not in the expected direction. The mean score decreased for same-trials with positive clue, implying that participants of Experiment 1 have favored “Different” responses for that clue. However, accuracy scores of negative trials remained at same low levels for both positive-clue and negative-clue trials, meaning that the participants did not

really detect the different pairs. It is possible that participants had implicitly associated the positive clue with different-trials but were biased toward "Same" responses for other clues.

The mean score drops of reversal and non-reversal groups were sufficiently different to imply that reversal-group participants incurred a performance cost. However, since no reliable indication was found that participants learned the roles of the clues before the reversal, this performance cost, at best, is an insufficient evidence of implicit learning.

Low accuracy scores indicate the test was too difficult to be used to understand the effects of implicit learning process. Especially the average scores of many participants in different-trials are barely above the chance level. Low accuracy scores might have resulted from fatigue and growing lack of attention because a typical session of Experiment 1 lasted quite long, about an hour or so.

Nevertheless, Experiment 1 served well as a pilot study that suggested the possibility of implicit learning in an auditory perception task involving probability relationships between structural elements of music. The participants' responses to the post-task questionnaires conclusively eliminated the possibility that any of the participants were performing the task by explicitly relying on the predictiveness relationships of the clues. Although some participants indicated that they were aware of the clues' presence and a few participants may have initially attributed some meanings to clues, no participant established a relationship between any clue and same or different-trials.

Chapter 5

EXPERIMENT 2

IX. OVERVIEW OF EXPERIMENT 2

A. Purpose

The purpose of Experiment 2 was to examine the process of implicit learning in auditory perception by utilizing the predictive context of melody clues embedded in musical melodies played at fast tempo.

B. Method

The main task in Experiment 2 was the same task commonly used throughout the present study: determining whether or not a pair of melodies was same or different. Melody clues, i.e., specific tone sequences embedded in melodies provided predictive information relating to whether or not a specific pair was likely to be a same-pair or a different-pair. Three different melody clues (see page 34) of three different predictive relationships with same-trials (75% for positive clue, 50% for negative clue and 25% for the neutral clue) were used in equal proportions throughout 150 trials that made up each experimental session. Experiment 2 employed the same 150 melodies used in Experiment 1, but the melodies were played at a fast tempo (200 beats per minute, a standard note lasts 1/200 of a

minute) so that participants would hear melodies more like complete patterns than a collection of individual notes. Each melody (including its clue) lasted about 4 seconds.

C. Participants

A total of 48 participated in this experiment, 29 in the reversal group, and 19 in the nonreversal group from various introductory psychology courses. The number of participants exceeding the 54% accuracy limit was 21 in the reversal group and 17 in the nonreversal group. For their participation, participants received two hours credit for a course requirement.

X. RESULTS OF EXPERIMENT 2

A. Analysis of Performance Changes before the 90th Trial

1. Mean Scores of Same-Trials

The following table shows the summary statistics for the accuracy scores of Experiment 2 for same-trials:

Table 5 Means, Standard Deviations and N for Accuracy Scores for Same-Trials before 90

Same-Trials				
Clue	Trial Block	Accuracy Score (Mean)	Accuracy Score (Std. Dev.)	N
Positive	1-45	0.7927	0.1712	38
	46-90	0.8173	0.1435	38
Negative	1-45	0.8241	0.2431	38
	46-90	0.7356	0.2843	38
Neutral	1-45	0.7984	0.1306	38
	46-90	0.7826	0.1368	38

The following figure shows the changes in the mean accuracy scores summarized in the above table:

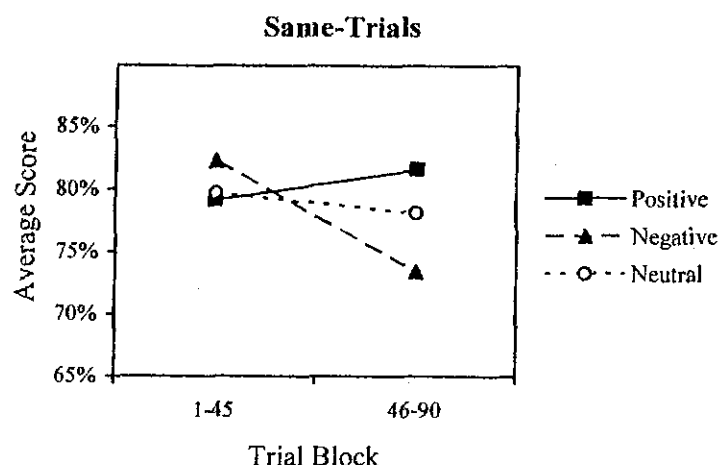


Figure 12 Mean Accuracy Scores of Same-Trials before 90

It is immediately noticeable that the accuracy scores in same-trials were fairly high. Furthermore, the accuracy scores of same-trials with all clues changed in the expected ways. Positive-same trials' mean score increased (mean = 0.7927 for trial block 1-45 and mean = 0.8173 for trial block 46-90), indicating that the participants of Experiment 2 perceived an association of the positive clue with same-trials. The mean score for negative clue decreased (mean = 0.8241 for trial block 1-45 and mean = 0.7356 for trial block 46-90); this was also expected, because having associated the negative clue with different trials, the participants would not have given the "Same" response to melody pairs with the negative clue.

A 3 (clue type) by 2 (trial block) repeated-measures ANOVA test verified that the changes in the accuracy scores were in the expected directions. However, there was no

significant main effect for clue type ($F(2,74) = 0.432$, $MSe = 2.802$, $p = 0.651$, $r = 0.10$) or for trial block ($F(1,37) = 1.119$, $MSe = 3.595$, $p = 0.297$, $r = 0.17$). No significant interaction was found between the clue type and the trial block either ($F(2,74) = 2.089$, $MSe = 2.989$, $p = 0.131$, $r = 0.22$). When neutral clue trials were eliminated from the analyses the contrast between negative and positive clue was found to be greater. Without neutral clue trials, the interaction was almost significant ($F(1,37) = 3.257$, $MSe = 3.733$, $p = 0.07$, $r = 0.28$). Further analysis revealed the hypothesized pattern of this interaction, that accuracy score for positive clue increased (0.7927 to 0.8173) from trial block (1-45) to trial block (46-90). However, this effect was not significant ($F(1,37) = 0.560$, $p = 0.459$, $r = 0.12$). Accuracy score for negative clue decreased from 0.8241 to 0.7356. However, this effect was also not significant ($F(1,37) = 2.516$, $p = 0.121$, $r = 0.25$). There was no significant main effect for clue type ($F(1,37) = 0.601$, $MSe = 3.993$, $p = 0.443$, $r = 0.12$). However, as hypothesized, participants did better with positive clue (0.805) overall than did with negative clue (0.780) and it is descriptive. There was no significant main effect for trial block ($F(1,37) = 0.915$, $MSe = 4.237$, $p = 0.345$, $r = 0.15$). Overall, participants in block 1-45 did better (0.808) than did participants in block 45-90 (0.776). However, this lack of a main effect is the result offsetting simple effects of trial block for the positive clue which showed an increase in average score from block1 to block2 as expected and of trial block for the negative clue which showed a decrease in average score from block1 to block2. Out of 38 participants, 25 participants' accuracy scores increased for positive clue. As for the negative clue, only 13 participants' accuracy scores increased.

2. Mean Scores of Different-Trials

The summary statistics for the accuracy scores for different-trials are shown in the following table:

Table 6 Means, Standard Deviations and N for Accuracy Scores for Different-Trials before 90

Different-Trials				
Clue	Trial Block	Accuracy Score (Mean)	Accuracy Score (Std. Dev.)	N
Positive	1-45	0.6918	0.2077	36
	46-90	0.6231	0.1912	36
Negative	1-45	0.6195	0.1859	36
	46-90	0.6587	0.1559	36
Neutral	1-45	0.5273	0.3152	36
	46-90	0.6079	0.2837	36

The following figure shows the changes in the mean scores listed in the above table:

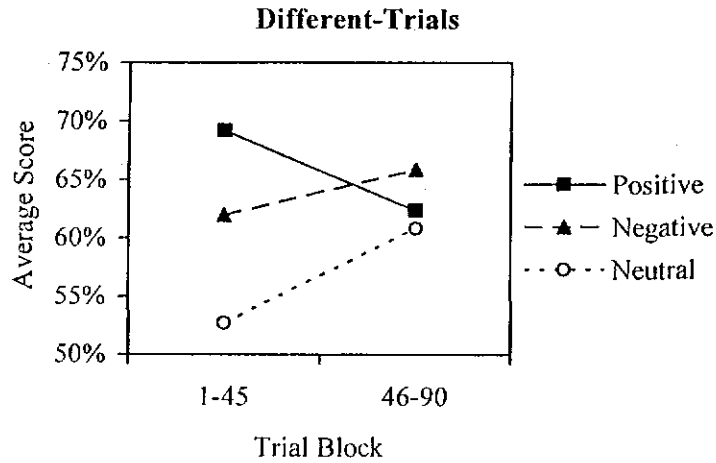


Figure 13 Mean Accuracy Scores of Different-Trials before 90

Mean scores of different-trials started out reasonably high for positive and negative clues, and they also changed in the expected directions. Mean score for the positive clue was 0.6918 for the first trial block (1-45) and 0.6231 for the second trial block (46-90). In other words, it did decrease as expected. Mean score for the negative clue was 0.6195 for the first trial block and 0.6587 for the second trial block; it increased as expected. According to a 3 (clue type) by 2 (trial block) repeated-measures ANOVA test, there is a significant main effect of clue type ($F(2,70) = 4.728$, $MSe = 3.4354$, $p = 0.012$, $r = 0.33$). There is, however, no significant main effect of trial block ($F(1,35) = 0.269$, $MSe = 5.828$, $p = 0.607$, $r = 0.08$). There is no significant interaction to be found between clue type and trial block ($F(2,70) = 2.112$, $MSe = 5.059$, $p = 0.129$, $r = 0.22$). When neutral clue trials were eliminated from the analyses, the contrast between negative and positive clue was

found to be greater. Without neutral clue trials, the interaction was almost significant ($F(1,37) = 3.338$, $MSe = 3.572$, $p = 0.07$, $r = 0.28$). Further analysis revealed the hypothesized pattern of this interaction, that accuracy score for positive clue decreased (0.6930 to 0.6331) from trial block (1-45) to trial block (46-90) as expected. However, this effect was not significant ($F(1,37) = 1.396$, $p = 0.245$, $r = 0.18$). Accuracy score for negative clue increased from 0.6170 to 0.6692 as expected. However, this effect was also not significant ($F(1,37) = 2.310$, $p = 0.137$, $r = 0.24$). There was no significant main effect for clue type ($F(1,37) = 0.840$, $MSe = 1.796$, $p = 0.365$, $r = 0.14$). Overall participants did better with positive clue (0.663) than the negative clue (0.643). There was no significant main effect for trial block ($F(1,37) = 0.016$, $MSe = 3.544$, $p = 0.900$, $r = 0.02$). Overall, participants in block 1-45 did better (0.655) than did participants in block 45-90 (0.651). However, this lack of a main effect is the result of offsetting simple effects of trial blocks for the positive clue which showed a decrease in average score from block1 to block2 and of trial block for the negative clue which showed an increase in average score from block1 to block2. Out of 38 participants, 22 participants' accuracy scores increased for negative clue. As for the positive clue, only 11 participants' accuracy scores increased.

3. Mean Scores of Positive-Clue Trials

When the mean accuracy scores of same-trials and different-trials of positive clue were compared, it was easier to conclude that the changes were in correct directions. The following figure shows the striking contrast between the changes in the mean scores of positive-same and positive-different trials.

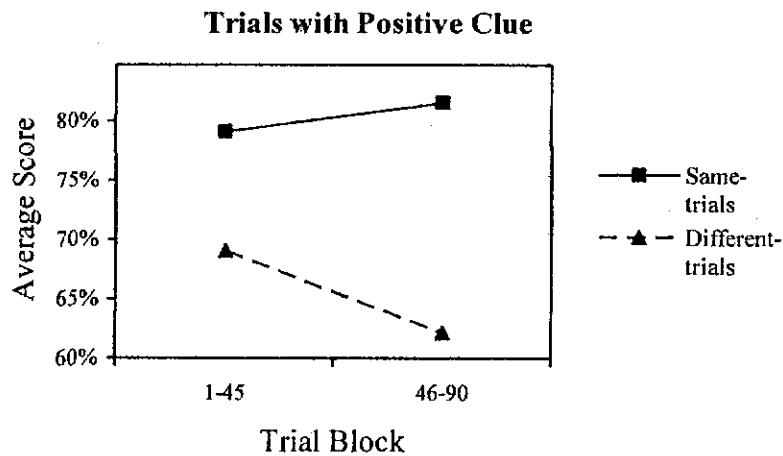


Figure 14 Mean Accuracy Scores of Positive Clue for Trials before 90

The repeated-measures ANOVA test showed no significant main effect of trial block for trials starting with the positive clue ($F(1,37) = 0.314$, $MSe = 3.760$, $p = 0.579$, $r = 0.08$). However, the significant main effect of trial type (i.e., same or different) was ascertained ($F(1,37) = 30.374$, $MSe = 2.522$, $p = 0.000$, $r = 0.67$). Overall, participants in same trial with positive clue did better (marginal mean = 0.805) than the participants in different trial (marginal mean = 0.780). No significant interaction between trial type and trial block ($F(1,37) = 12.137$, $MSe = 3.173$, $p = 0.152$, $r = 0.23$).

4. Mean Scores of Negative-Clue Trials

When the accuracy scores of same-trials and different-trials of negative clue were compared, the following figure was obtained:

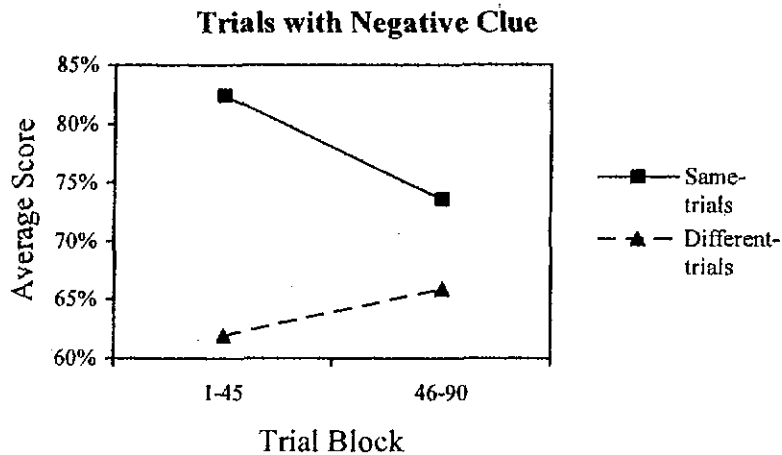


Figure 15 Mean Accuracy Scores of Negative Clue for Trials before 90

The repeated-measures ANOVA test shows no significant main effect of trial block for trials starting with the negative clue ($F(1,37) = 0.353$, $MSe = 3.552$, $p = 0.556$, $r = 0.3$). However, this lack of a main effect is again the result of offsetting simple effects of trial blocks for both same and different trials. However, there was a significant main effect of trial type (i.e., same or different) ($F(1,37) = 12.439$, $MSe = 5.714$, $p = 0.001$, $r = 0.50$). As hypothesized, accuracy score were lower (mean= 0.643) for same trial for negative clue than for the different trials (mean= 0.780). No significant interaction existed between trial type and trial block ($F(1,37) = 4.086$, $MSe = 4.601$, $p = 0.051$, $r = 0.31$).

B. Analysis of Performance Changes after the 90th Trial

1. Score Drops In the Reversal and Nonreversal groups

As it was done for Experiment 1, the performance cost of the reversal was analyzed by looking at the change in the combined mean score of trials with positive and negative clues from trial block 2 (46-90) to trial block 3 (91-120). The following table lists the combined mean scores of same-trials and different-trials for those two blocks:

Table 7 Means, Standard Deviations and N for Accuracy Scores
of Reversal and Nonreversal groups for Trials before and after 90

Participant Group	Trial Blocks	
	46-90	91-120
Reversal group	Mean = 0.6931	Mean = 0.6829
	Std. Dev. = 0.2140	Std. Dev. = 0.2841
	N = 84	N = 84
Nonreversal group	Mean = 0.7393	Mean = 0.6643
	Std. Dev. = 0.2117	Std. Dev. = 0.2430
	N = 68	N = 68

The changes were not found to be significant for the reversal group ($t(83) = -0.284$, $p > 0.05$) but they were significant for the nonreversal group ($t(67) = -2.073$, $p < 0.05$). In other words, contrary to the expectations, the performance drop in nonreversal group was greater than the one observed in the experimental group. More detailed analyses revealed that this effect was due to unexpectedly fast re-learning of clues in the reversal group - same-trials after the reversal of the rule. Mean score of same-trials with positive clue was 0.8058 before the reversal (for trial block 46-90) and 0.7352 after the reversal (for trial

block 91-120). In other words, it decreased as expected (see Table 8 on page 70).

Mean score for same-trials with negative clue was 0.7041 before the reversal and 0.7706 after the reversal; it showed an increase as expected. In summary, the mean scores for same-trials did show the changes expected from participants who implicitly learned the use the clues. A 2 by 2 repeated measures ANOVA was applied to trials with the positive clue in order to see if this effect was significant. A significant interaction was found between trial block and trial type ($F(1,20) = 0.224$, $MSe = 4.293$, $p < 0.05$).

2. Mean Scores of Trials after 90

a) Mean Scores of Same-Trials in the Reversal group

The following table summarizes the summary statistics for same-trials of reversal group:

Table 8 Means, Standard Deviations and N for Accuracy Scores
of Reversal group for Same-Trials after 90

Reversal group: Same-Trials				
Clue	Trial Block	Accuracy Score (Mean)	Accuracy Score (Std. Dev.)	N ¹
Positive	46-90	0.8058	0.1485	21
	91-120	0.7352	0.16711	20
	121-150	0.8407	0.2339	20
Negative	46-90	0.7041	0.2692	21
	91-120	0.7706	0.1504	20
	121-150	0.7965	0.2069	20
Neutral	46-90	0.7897	0.1266	21
	91-120	0.6967	0.3537	20
	121-150	0.8217	0.2290	20

¹ The N values do not always match; because the clue type and the trial type was randomly determined for every trial of every session, trial type-clue type combinations were not equally distributed over trial blocks.

The following figure shows the changes in the mean accuracy scores summarized in the above table:

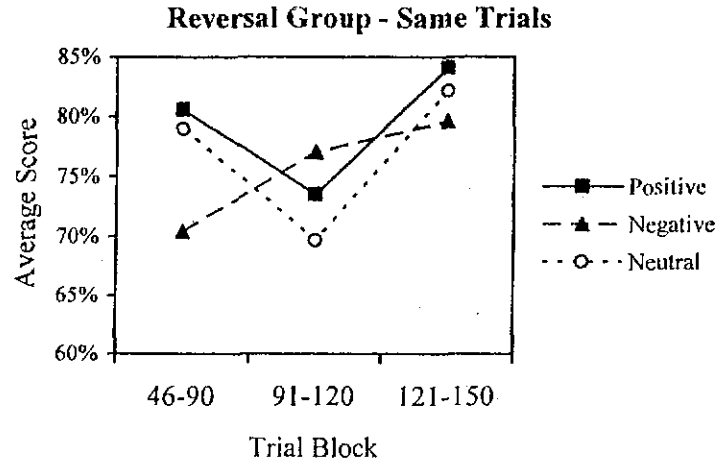


Figure 16 Mean Accuracy Scores of Reversal group for Same-Trials after 90

The repeated-measures ANOVA test of same-trials in reversal group after 90 did not show a significant main effect of clue type ($F(2,38) = 0.224$, $MSe = 4.293$, $p = 0.800$, $r = 0.10$). The main effect of trial block, on the other hand, was significant ($F(1,19) = 5.190$, $MSe = 4.224$, $p < 0.05$, $r = 0.46$). The interaction between clue type and trial block, however, was not significant ($F(2,38) = 0.513$, $MSe = 5.363$, $p = 0.603$, $r = 0.16$). An evaluation of these results showed that the mean accuracy score of negative clue continued to increase as expected (mean = 0.7706 for trial block 91-120 and mean = 0.7965 for trial block 121-150). However, it seems that clue learning deteriorated for the positive clue (mean = 0.7352 for trial block 91-120 and mean = 0.8407 for trial block 121-150). When neutral clue trials were eliminated from the analyses, interaction was still not significant ($F(1,19) = 1.035$, $p = 0.321$, $MSe = 2.153$, $r = 0.22$). There was no significant main effect of

trial block $F(1,19) = 2.343$, $p = 0.141$, $MSe = 4.129$, $r = 0.32$). Overall, accuracy score was lower (0.751) for trial block 90-120 than for the trial block 121-150 (0.819). There was no significant main effect of clue type $F(1,19) = 0.350$, $p = 0.561$, $MSe = 3.868$, $r = 0.13$). Overall, accuracy score was higher for positive clue trials (0.798) than for the negative clue trials (0.773).

b) Mean Scores of Different-Trials in the Reversal group

The following table summarizes the summary statistics for different-trials of the reversal group:

Table 9 Means, Standard Deviations and N for Accuracy Scores
of Reversal group for Different-Trials after 90

Reversal group: Different-Trials				
Clue	Trial Block	Accuracy Score (Mean)	Accuracy Score (Std. Dev.)	N ²
Positive	46-90	0.5928	0.2127	19
	91-120	0.5961	0.2246	21
	121-150	0.6217	0.3327	21
Negative	46-90	0.6270	0.1420	19
	91-120	0.6071	0.3847	21
	121-150	0.6984	0.3359	21
Neutral	46-90	0.5325	0.2670	19
	91-120	0.4912	0.2380	21
	121-150	0.5089	0.2069	21

² The N values do not always match; because the clue type and the trial type was randomly determined for every trial of every session, trial type-clue type combinations were not equally distributed over trial blocks.

The following figure shows the changes in the mean accuracy scores summarized in the above table:

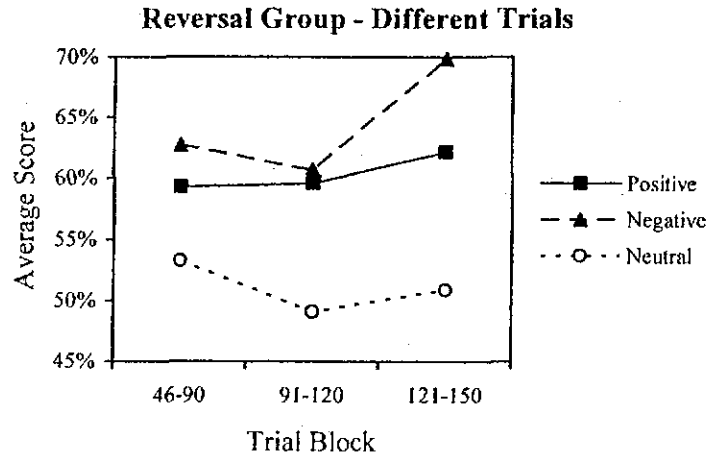


Figure 17 Mean Accuracy Scores of Reversal group for Different-Trials after 90

The repeated-measures ANOVA test of different-trials in the reversal group after 90 did show a significant main effect of clue type ($F(2,40) = 3.623$, $MSe = 7.170$, $p = 0.036$, $r = 0.39$). The main effect of trial block was not significant. ($F(1,20) = 1.067$, $MSe = 5.930$, $p = 0.314$, $r = 0.22$). The interaction between clue type and trial block was not significant ($F(2,40) = 0.277$, $MSe = 6.190$, $p = 0.760$, $r = 0.11$). However, an evaluation of these results showed that learning of the negative clue deteriorated after the 90th trial. Mean accuracy score of trials with negative clue increased instead of decreasing towards the end (mean = 0.6071 for trial block 91-120 and mean = 0.6984 for trial block 121-150). However, for the trials with positive clue, the mean score increased as expected (mean = 0.5961 for trial block 91-120 and mean = 0.6217 for trial block 121-150). When neutral clue trials were eliminated from the analyses, interaction was still not significant ($F(1,19) =$

0.384, $p = 0.543$, $MSe = 5.918$, $r = 0.13$). There was no significant main effect of trial block $F(1,19) = 0.858$, $p = 0.365$, $MSe = 8.341$, $r = 0.2$). Overall, accuracy score was lower (0.602) for trial block 90-120 than for the trial block 121-150 (0.660). There was no significant main effect of clue type $F(1,19) = 0.378$, $p = 0.546$, $MSe = 0.107$, $r = 0.13$). Overall, accuracy score was lower for positive clue (0.609) than for the negative clue (0.653).

c) Mean Scores of Same-Trials in the Nonreversal group

The same analysis was also applied to the nonreversal group scores. The following table presents the summary statistics for same-trials of the nonreversal group of Experiment 2:

Table 10 Means, Standard Deviations and N for Accuracy Scores
of Nonreversal group for Same-Trials after 90

Nonreversal group: Same-Trials				
Clue	Trial Block	Accuracy Score (Mean)	Accuracy Score (Std. Dev.)	N ³
Positive	46-90	0.8315	0.1402	17
	91-120	0.8010	0.2168	16
	121-150	0.8461	0.1890	16
Negative	46-90	0.7745	0.3058	17
	91-120	0.7440	0.2292	16
	121-150	0.5677	0.3918	16
Neutral	46-90	0.7738	0.1518	17
	91-120	0.7647	0.1902	16
	121-150	0.7485	0.2315	16

³ The N values do not always match; because the clue type and the trial type was randomly determined for every trial of every session, trial type-clue type combinations were not equally distributed over trial blocks.

The following figure shows the changes in the mean scores listed in the above table:

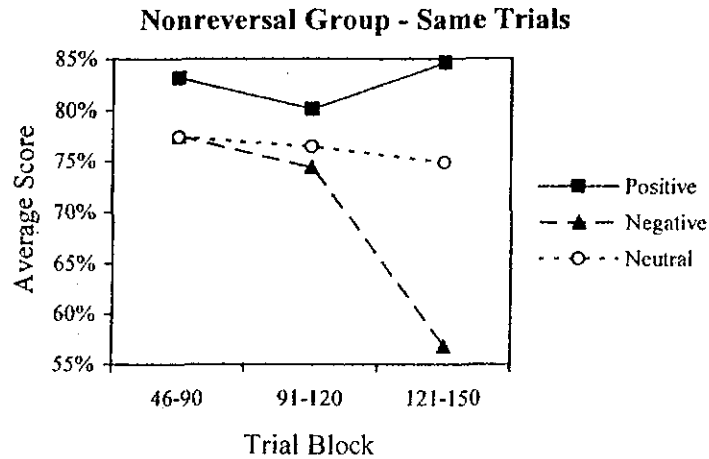


Figure 18 Mean Accuracy Scores of Nonreversal group for Same Trials after 90

The repeated-measures ANOVA test of same-trials in nonreversal group after 90 did show a significant main effect of clue type ($F(2,30) = 5.748$, $MSe = 3.968$, $p = 0.008$, $r = 0.52$). The main effect of trial block was not significant ($F(1,15) = 0.775$, $MSe = 7.480$, $p = 0.392$, $r = 0.22$). The interaction between clue type and trial block was not significant ($F(2,30) = 1.497$, $MSe = 6.988$, $p = 0.240$, $r = 0.30$). However, an evaluation of these results showed that the mean accuracy score of trials with negative clue continued to decrease as expected (mean = 0.7440 for trial block 91-120 and mean = 0.5677 for trial block 121-150). The mean score for trials with positive clue also fulfilled the expectations by increasing from Block 3 to Block 4 (mean = 0.8010 for trial block 91-120 and mean = 0.8461 for trial block 121-150). When neutral clue trials were eliminated from the analyses, interaction was still not significant ($F(1,15) = 2.402$, $p = 0.142$, $MSe = 8.165$, $r = 0.37$).

However, there was still a significant main effect of clue type $F(1,15) = 9.158$, $p = 0.009$, $MSe = 4.914$, $r = 0.61$). Overall, accuracy score was higher (0.824) for positive clue than for the negative clue (0.656). There was no significant main effect of trial block $F(1,15) = 0.760$, $p = 0.397$, $MSe = 9.061$, $r = 0.21$). Overall, accuracy score was higher for the trial block 90-120 (0.773) than for the trial block 121-150 (0.707).

A paired-samples t-test was also applied to determine if mean scores of negative and positive clues were significantly different from each other. It was found that negative and positive clues were significantly different from each other ($t(16) = 2.527$, $p < 0.05$).

d) Mean Scores of Different-Trials in the Nonreversal group

The following table presents the summary statistics for different-trials of the nonreversal group:

Table 11 Means, Standard Deviations and N for Accuracy Scores
of Nonreversal group for Different-Trials after 90

Nonreversal group: Different-Trials				
Clue	Trial Block	Accuracy Score (Mean)	Accuracy Score (Std. Dev.)	N
Positive	46-90	0.6570	0.20331	17
	91-120	0.5802	0.3138	17
	121-150	0.5096	0.2291	17
Negative	45-90	0.6942	0.16727	17
	91-120	0.5875	0.1911	17
	121-150	0.6003	0.1919	17
Neutral	45-90	0.6922	0.28560	17
	91-120	0.4272	0.4165	17
	121-150	0.4265	0.2855	17

The following figure shows the changes in the mean score listed in the above table:

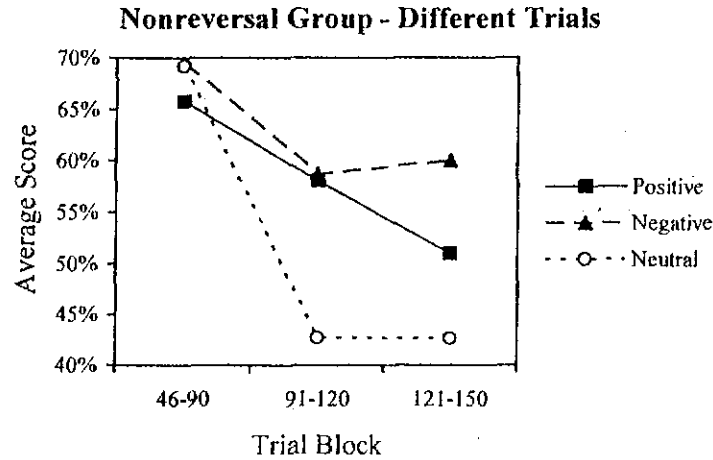


Figure 19 Mean Accuracy Scores of Nonreversal group for Different-Trials after 90

The repeated-measures ANOVA test of different-trials in the nonreversal group after 90 did not show a significant main effect of clue type ($F(2,32) = 2.905$, $MSe = 9.688$, $p = 0.069$, $r = 0.39$). The main effect of trial block was not significant either ($F(1,16) = 0.169$, $MSe = 5.734$, $p > 0.687$, $r = 0.1$). The interaction between clue type and trial block was not significant ($F(2,32) = 0.266$, $MSe = 6.404$, $p = 0.768$, $r = 0.12$). However, an evaluation of these results showed that accuracy score of negative clue continued to increase as expected (mean = 0.5875 for trial blocks 91-120 and mean = 0.6003 for trial blocks 121-150). Accuracy score of the positive clue also continued to decrease as expected (mean = 0.5802 for trial blocks 91-120 and mean = 0.5096 for trial blocks 121-150). When neutral clue trials were eliminated from the analyses, interaction was still not significant ($F(1,16) = 1.146$, $p = 0.300$, $MSe = 2.581$, $r = 0.25$). There was no significant

main effect of trial block $F(1,16) = 0.310$, $p = 0.585$, $MSe = 4.518$, $r = 0.13$).

Overall, accuracy score was higher (0.584) for trial block 90-120 than for the trail block 121-150 (0.555). There was no significant main effect of clue type $F(1,16) = 0.953$, $p = 0.344$, $MSe = 4.281$, $r = 0.23$). Overall, accuracy score was lower for positive clue (marginal mean = 0.545) than for the negative clue (marginal mean = 0.594).

C. Summary of the Results

The changes in the task performance of the participants of experiment 2 until the 90th trial show a clear evidence of clue learning. Mean accuracy scores were generally high, an average of 80% for same-trials and about 70% for different-trials. Positive-same trials' mean score increased (mean = 0.7927 for trial block 1-45 and mean = 0.8173 for trial block 46-90), indicating that the participants of Experiment 2 perceived the strong association of the positive clue with same-trials. The mean score for negative clue trials decreased (mean = 0.8241 for trial block 1-45 and mean = 0.7356 for trial block 46-90); this was also expected, because having associated the negative clue with different trials, the participants would not have given the "Same" response to melody pairs with the negative clue.

As for different trials, mean score for the positive clue was 0.6918 for the first trial block (1-45) and 0.6231 for the second trial block (46-90). In other words, it did decrease as expected. Mean score for the negative clue was 0.6195 for the first trial block and 0.6587 for the second trial block; it increased as expected.

When we compare the mean scores of reversal group and nonreversal group we also see the evidence of implicit learning of the rules. Nonreversal group scores after 90th

trial continued to fulfill our expectations. The results showed that the mean accuracy score of trials with negative clues continued to decrease as expected in same trials. The mean score for trials with positive clue continued to increase as expected in same trials. As for the different trials, the mean accuracy score of negative clue also continued to increase as expected. The mean accuracy score of the positive clue continued to decrease again as expected. The reversal group scores after the 90 th trial does not show an expected score drop however, the lack of significant score drop in the reversal group can be attributed to a quick re- learning of the new predictiveness relationship in the same trials.

When the accuracy scores of same-trials and different-trials of negative clue were compared, it was clear that participants, in general, better understood the association between the negative clue and the trial type than the association between the positive clue and the trial type.

As in Experiment 1, the participants' responses to the post-task questionnaires conclusively eliminated the possibility of explicit task performance. Only some participants' responses indicated that they considered the possibility that clues had some meanings, but no participant discovered what those meanings could be.

Chapter 6

THE METHODOLOGY AND RESULTS OF EXPERIMENT 3

XI. OVERVIEW OF EXPERIMENT 3

A. Purpose

The purpose of Experiment 3 was to examine the process of implicit learning in auditory perception by utilizing the predictive context of rhythm clues embedded in musical melodies.

B. Method

The main task in Experiment 3 was the same task commonly used throughout the present study: determining whether or not a pair of melodies was same or different. Rhythm clues, i.e., sequences of notes with unchanging lengths were embedded in melodies provided predictive information relating to whether or not a specific pair was likely to be a same-pair or a different-pair. Three different rhythm clues (see page 35) of three different predictive relationships with same-trials (75% for positive clue, 50% for neutral clue and 25% for negative clue) were used in equal proportions throughout 150 trials that made up each experimental session. Experiment 3 employed the same 150 melodies used in the other experiments, but as in Experiment 1, the melodies were played at a slow tempo (120 beats per minute, a standard note lasts for 1/120 of a minute) to

enable the participants to better distinguish the individual notes. Each melody (including its clue) lasted about 7 seconds.

C. Participants

A total of 43 participated in this experiment, 23 in the reversal group, and 20 in the nonreversal group from various introductory psychology courses. Among those participants, 23 participants of the reversal group and 16 participants of the nonreversal group performed beyond the 54% limit. For their participation, participants received two hours credit for a course requirement.

XII. RESULTS OF EXPERIMENT 3

A. Analysis of Performance Changes before the 90th Trial

1. Mean Scores of Same-Trials

The following table shows the summary statistics for the accuracy scores of Experiment 3 for same-trials:

Table 12 Means, Standard Deviations and N for Accuracy Scores for Same-Trials before 90

Same-Trials				
Clue	Trial Block	Accuracy Score (Mean)	Accuracy Score (Std. Dev.)	N
Positive	1-45	0.7378	0.1461	39
	46-90	0.7711	0.1928	39
Negative	1-45	0.8060	0.24467	39
	46-90	0.6693	0.2837	39
Neutral	1-45	0.7295	0.1471	39
	46-90	0.6898	0.1603	39

The following figure shows the changes in the mean accuracy scores summarized in the above table:

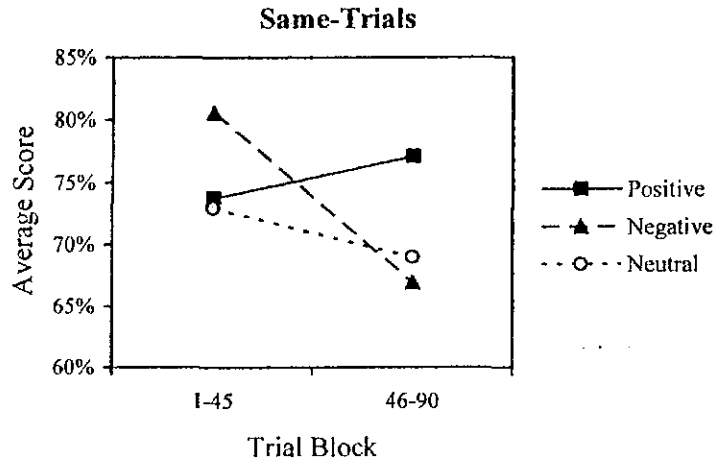


Figure 20 Mean Accuracy Scores of Same-Trials before 90

According to 3 (clue type) by 2 (trial block) repeated-measures ANOVA test there is no significant main effect of clue type ($F(2,76) = 1.087$, $MSe=3.673$, $p = 0.342$, $r = 0.16$), not a significant main effect is found for trial block ($F(1,38) = 3.625$, $MSe = 3.671$, $p = 0.065$), $r = 0.29$. The interaction between clue type and the trial block was found to be significant ($F(2,76) = 4.597$, $MSe = 3.069$, $p = 0.013$, $r = 0.32$), in the expected direction. When neutral clue trials were eliminated from the analyses, the contrast between negative and positive clue was again found to be greater. Without neutral clue trials, the interaction was still significant ($F(1,38) = 6.318$, $MSe= 4.457$, $p = 0.016$, $r= 0.37$). Further analysis revealed the hypothesized pattern of this interaction, that accuracy score for positive clue increased (0.7378 to 0.7711) from trial block (1-45) to trial block (46-90). However, this effect was not significant ($F(1,38)=0.997$, $p = 0.324$, $r = 0.16$). Accuracy score for negative

clue decreased from 0.8060 to 0.6693. This effect was significant ($F(1,38)=5.676$, $p = 0.022$, $r = 0.36$). There was no significant main effect for clue type ($F(1,38) = 0.245$, $MSe = 4.476$, $p = 0.623$, $r = 0.07$). However, as hypothesized, participants did better with positive clue (0.754) overall than did negative clue (0.738). There was no significant main effect for trial block ($F(1,38) = 2.525$, $MSe = 4.127$, $p = 0.120$, $r = 0.24$). However, this lack of a main effect is the result offsetting simple effects of trial block for the positive clue which showed an increase in average score from block1 to block2 as expected and of trial block for the negative clue which showed a decrease in average score from block1 to block2. Out of 39 participants, 24 participants' accuracy scores increased for positive clue. As for the negative clue, only 14 participants' accuracy scores increased.

An evaluation of these results showed that the expected performance improvements are observed in same-trials. Firstly, the mean accuracy score for positive clue increased as expected (mean = 0.7378 for trial block 1-45 and mean = 0.7711 for trial block 46-90). Secondly, the average score for negative clue decreased (mean = 0.8060 for trial block 1-45 and mean = 0.6693 for trial block 46-90).

2. Mean Scores of Different-Trials

The summary statistics for the accuracy scores for different-trials are shown in the following table:

Table 13 Means, Standard Deviations and N for Accuracy Scores for Different-Trials before 90

Different-Trials				
Clue	Trial Block	Accuracy Score (Mean)	Accuracy Score (Std. Dev.)	N
Positive	1-45	0.5141	0.2135	38
	46-90	0.5005	0.2029	38
Negative	1-45	0.5158	0.2139	38
	46-90	0.5388	0.1747	38
Neutral	1-45	0.5004	0.3237	38
	46-90	0.5877	0.3459	38

The following figure shows the changes in the mean scores listed in the above table:

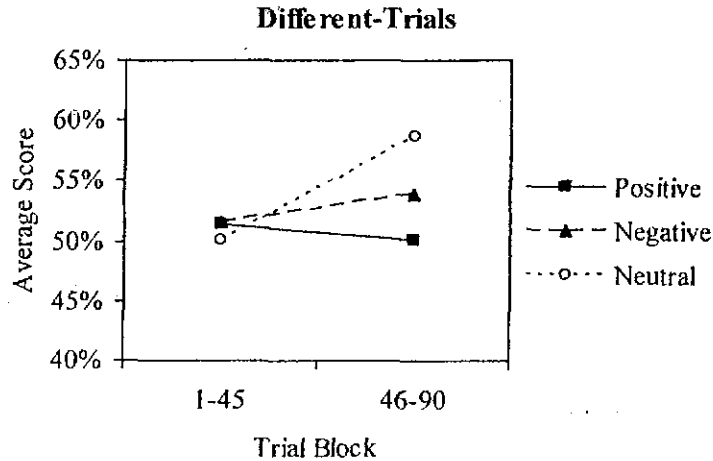


Figure 21 Mean Accuracy Scores of Different-Trials before 90

According to a 3 (clue type) by 2 (trial block) repeated-measures ANOVA test, there was not a significant main effect of clue type ($F(2,74) = 0.582$, $MSe = 4.419$, $p = 0.561$, $r = 0.12$). Neither was there a significant main effect of trial block ($F(1,37) = 1.379$, $MSe = 4.296$, $p = 0.248$, $r = 0.18$). Finally, there was no significant interaction to be found between clue type and trial block ($F(2,74) = 0.926$, $MSe = 5.350$, $p = 0.401$, $r = 0.15$). When neutral clue trials were eliminated from the analyses, the interaction was still not significant ($F(1,38) = 0.521$, $MSe = 3.481$, $p = 0.475$, $r = 0.11$). There was no significant main effect for clue type ($F(1,38) = 0.665$, $MSe = 2.592$, $p = 0.420$, $r = 0.13$). However, as hypothesized, overall participants did better with negative clue (0.530) than did positive clue (0.509). There was no significant main effect for trial block ($F(1,37) = 0.915$, $MSe = 4.237$, $p = 0.345$, $r = 0.15$). Overall, participants in block 1-45 (marginal mean = 0.520) and

block 45-90 (marginal mean= 0.520) did equally well. Out of 39 participants, 28 participants' accuracy scores increased for negative clue. As for the positive clue, 12 participants' accuracy scores increased.

However, an evaluation of cell means showed that the expected performance improvements were observed for different-trials. Mean score for positive clue was 0.5141 for the first trial block (1-45) and 0.5005 for the second trial block (46-90). However, the mean scores of those blocks were not much different from each other. Mean score for negative clue was 0.5158 for the first trial block and 0.5388 for the second trial block, an increase as expected, but an insignificant one.

3. Mean Scores of Positive-Clue Trials

In order to better focus on the changes occurring in the scores for positive clue, accuracy scores of same-trials and different-trials of positive clue were compared. As the following figure shows, the changes were in expected directions and in contrast with each other, but the contrast is a weak one.

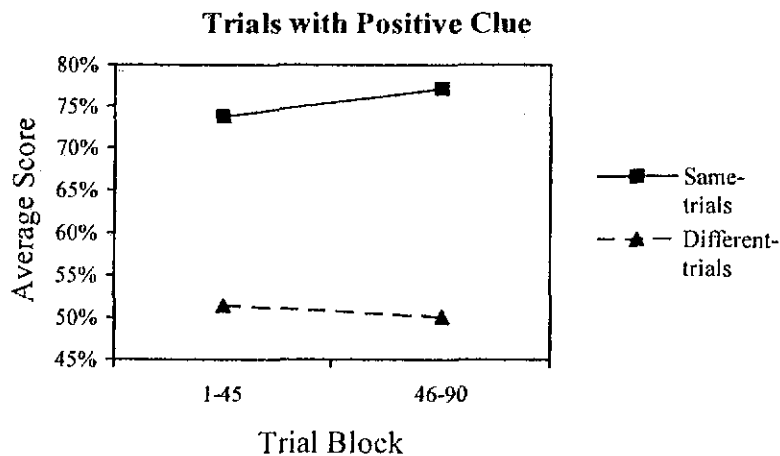


Figure 22 Mean Accuracy Scores of Positive Clue for Trials before 90

The repeated-measures ANOVA test showed no significant main effect of trial block for trials starting with the positive clue ($F(1,38) = 0.049$, $MSe = 2.776$, $p = 0.826$, $r = 0.03$). There was a significant main effect of trial type (i.e., same or different) ($F(1,38) = 54.149$, $MSe = 4.323$, $p = 0.000$, $r = 0.75$). As hypothesized, accuracy score was higher for the same trials (mean= 0.754) with positive clue than the accuracy score for the different trials (mean= 0.509). However, no significant interaction between trial type and trial block was found ($F(1,38) = 0.785$, $MSe = 3.726$, $p = 0.381$, $r = 0.14$).

4. Mean Scores of Negative-Clue Trials

When the accuracy scores of same-trials and different-trials of negative clue were compared, the following figure was obtained:

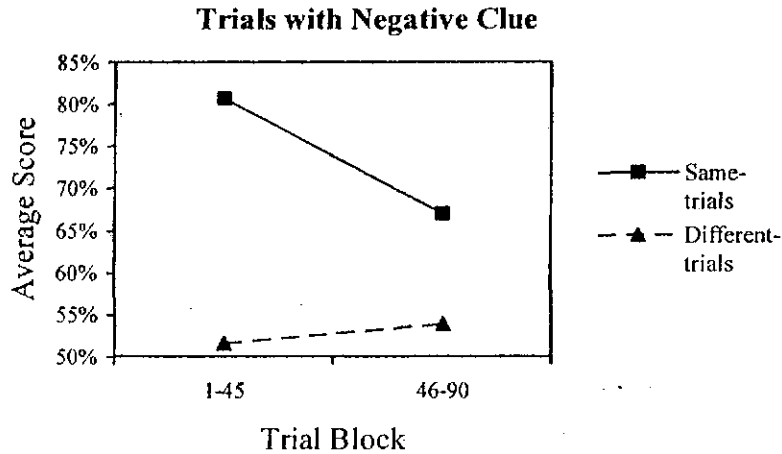


Figure 23 Mean Accuracy Scores of Negative Clue for Trials before 90

The repeated-measures ANOVA test did not find a significant main effect of trial block for trials starting with the negative clue ($F(1,38) = 2.804$, $MSe = 4.600$, $p = 0.102$, $r = 0.26$). The main effect of trial type (i.e., same or different) was significant ($F(1,38) = 21.044$, $MSe = 7.955$, $p = 0.000$, $r = 0.59$) and so was the interaction between trial type and trial block ($F(1,38) = 6.664$, $MSe = 3.667$, $p = 0.014$, $r = 0.38$). Further analysis revealed the hypothesized pattern of this interaction, that for same trials accuracy score was higher for the first trial block (1-45) than the second trial block (46-90) $F(1,38) = 5.676$, $p = 0.022$, $r = 0.36$. However, for the different trial, two types of trials worked equally well $F(1,38) = 0.494$, $p = 0.486$, $r = 0.36$).

B. Analysis of Performance Changes after the 90th Trial

1. Score Drops in the Reversal and Nonreversal groups

The performance cost of the reversal is apparent in the substantial drops in the mean scores for all clues in both participant groups. In order to determine the significance of the score drops observed for positive and negative clues, the mean scores of those clues were analyzed together for the trial blocks 46-90 and 91-120. The following table lists the combined mean scores of same-trials and different-trials for those two blocks:

Table 14 Means, Standard Deviations and N for Accuracy Scores
of Experiment and Nonreversal groups for Trials before and after 90

Participant Group	Trial Blocks	
	46-90	91-120
Reversal group	Mean = 0.6334	Mean = 0.5941
	Std. Dev. = 0.2243	Std. Dev. = 0.3044
	N = 92	N = 92
Nonreversal group	Mean = 0.6010	Mean = 0.5446
	Std. Dev. = 0.2919	Std. Dev. = 0.2721
	N = 64	N = 64

The mean accuracy scores dropped both in the reversal group and in the group, but the score drops were not found to be significant ($t(91) = -0.996$, $p > 0.05$ for the reversal group and $t(63) = -1.180$, $p > 0.05$ for the nonreversal group). The mean score drop for the reversal group was less than the mean score drop experienced by nonreversal-group participants. This was contrary to the expectations and implied that either implicit learning did not occur in reversal group or the score drops were caused by factors other than the

reversal of the predictiveness relationships. However, more detailed analyses showed that the unexpected relearning pattern after the reversal of the rule (see Table 21 and Figure 29) in different trials was probably the reason why the score drop for the reversal group was weaker than expected. The mean score for negative clue in different trial was 0.5556 for before reversal of the rule and 0.4732 for after the reversal. The mean score for positive clue in different trial increased unexpectedly (mean = 0.5002 before the reversal, for trial block 46-90, and mean = 0.5860 for after the reversal, for trial block 91-120). The most likely reason is that the reversal-group participants immediately re-learned the reversed predictiveness relationships of the positive clue. A 2 by 2 repeated measures ANOVA test was conducted on trials with positive clue in order to determine how significant that effect was. The test showed a significant interaction between trial block and trial type ($F(1,20) = 0.224$, $MSe = 4.293$, $p < 0.05$).

2. Mean Scores of Trials after 90

a) Mean Scores of Same-Trials in the Reversal group

The following table summarizes the summary statistics for same-trials of reversal group:

Table 15 Means, Standard Deviations and N for Accuracy Scores
of Reversal group for Same-Trials after 90

Reversal group: Same-Trials				
Clue	Trial Block	Accuracy Score (Mean)	Accuracy Score (Std. Dev.)	N ⁴
Positive	46-90	0.7849	0.1834	23
	91-120	0.7501	0.2322	22
	121-150	0.7258	0.2593	22
Negative	46-90	0.6929	0.2221	23
	91-120	0.6063	0.2266	22
	121-150	0.6945	0.1773	22
Neutral	46-90	0.7160	0.1590	23
	91-120	0.6891	0.2748	22
	121-150	0.6653	0.3326	22

⁴ The N values do not always match; because the clue type and the trial type was randomly determined for every trial of every session, trial type-clue type combinations were not equally distributed over trial blocks.

The following figure shows the changes in the mean accuracy scores summarized in the above table:

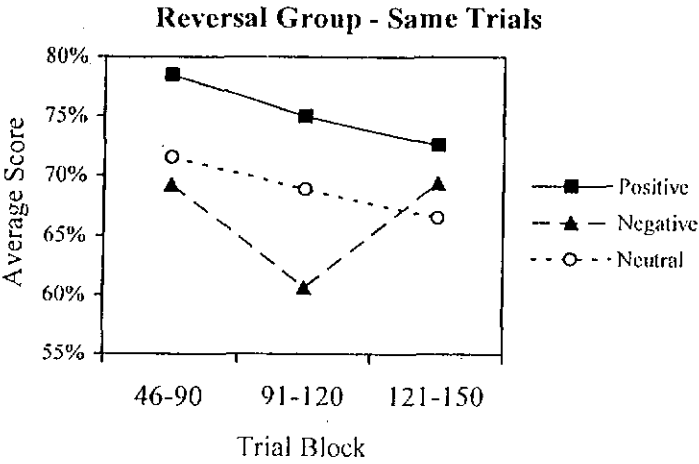


Figure 24 Mean Accuracy Scores of Reversal group for Same-Trials after 90

The repeated-measures ANOVA test of same-trials in reversal group after 90 did not show a significant main effect of clue type ($F(2,42) = 1.695$, $MSe = 5.225$, $p = 0.196$, $r = 0.27$). The main effect of trial block was not significant ($F(1,22) = 0.99$, $MSe = 5.942$, $p = 0.756$, $r = 0.07$). The interaction between clue type and trial block was not significant ($F(2,42) = 0.748$, $MSe = 6.179$, $p = 0.480$, $r = 0.18$). However, it should be noted that the mean accuracy score of trials with negative clue continued to increase as expected (mean = 0.6063 for trial block 91-120 and mean = 0.6945 for trial block 121-150). The mean accuracy score of the positive clue continued to decrease, and that was also expected (mean = 0.7501 for trial block 91-120 and mean = 0.7258 for trial block 121-150). When neutral clue trials were eliminated from the analyses, interaction was still not significant ($F(1,22) = 1.287$, $p = 0.269$, $MSe = 7.589$, $r = 0.23$). There was no significant main effect of trial block

$F(1,22) = 0.630$, $p = 0.436$, $MSe = 4.017$, $r = 0.16$). Overall, accuracy score was lower (0.663) for trial block 90-120 than for the trial block 121-150 (0.696). However, there was a significant main effect of clue type $F(1,19) = 5.337$, $p = 0.031$, $MSe = 2.687$, $r = 0.44$). Overall, accuracy score was higher for positive clue (0.719) than for the negative clue (0.640).

b) Mean Scores of Different-Trials in the Reversal group

The following table summarizes the summary statistics for different-trials of the reversal group:

Table 16 Means, Standard Deviations and N for Accuracy Scores
of Reversal group for Different-Trials after 90

Reversal group: Different-Trials				
Clue	Trial Block	Accuracy Score (Mean)	Accuracy Score (Std. Dev.)	N
Positive	46-90	0.5002	0.1954	23
	91-120	0.5860	0.2411	22
	121-150	0.4660	0.2170	22
Negative	45-90	0.5556	0.1854	23
	91-120	0.4732	0.4266	22
	121-150	0.4659	0.4031	22
Neutral	45-90	0.6442	0.3343	23
	91-120	0.5011	0.2164	22
	121-150	0.4609	0.2090	22

The following figure shows the changes in the mean accuracy scores summarized in the above table:

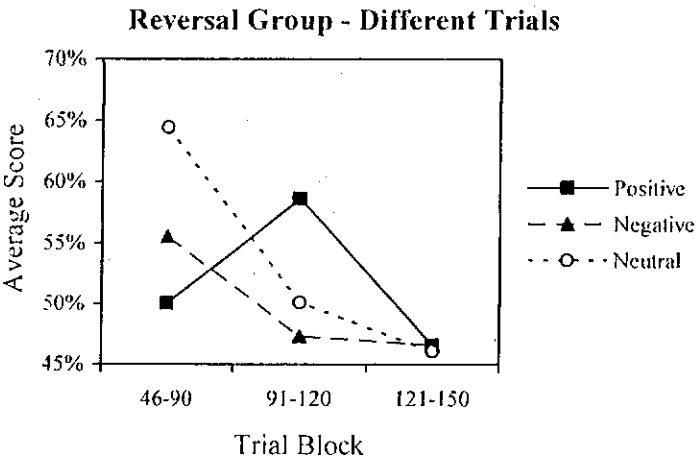


Figure 25 Mean Accuracy Scores of Reversal group for Different-Trials after 90

The repeated-measures ANOVA test of different-trials in the reversal group after 90 did not show a significant main effect of clue type ($F(2,42) = 0.550$, $MSe = 7.129$, $p = 0.581$, $r = 0.16$). The main effect of trial block was not significant either ($F(1,21) = 1.315$, $MSe = 7.824$, $p = 0.264$, $r = 0.24$). The interaction between clue type and trial block was not significant ($F(2,42) = 0.551$, $MSe = 6.708$, $p = 0.580$, $r = 0.16$). An evaluation of these results show that accuracy score of negative clue remained about the same (mean = 0.4732 for trial block 91-120 and mean = 0.4659 for trial block 121-150). However, accuracy score of the positive clue decreased (mean = 0.5860 for trial block 91-120 and mean = 0.4660 for trial block 121-150). In short, no evidence of re-learning was observed for either the positive or the negative clue. When neutral clue trials were eliminated from the analyses, interaction was still not significant ($F(1,21) = 0.648$, $p = 0.430$, $MSe = 0.108$, $r = 0.17$).

There was no significant main effect of trial block $F(1,21) = 1.186$, $p = 0.288$, $MSe = 7.513$, $r = 0.23$). Overall, accuracy score was higher (0.530) for trial block 90-120 than for the trial block 121-150 (0.466). There was no significant main effect of clue type $F(1,21) = 0.884$, $p = 0.358$, $MSe = 7.937$, $r = 0.2$). Overall, accuracy score was higher for positive clue (0.526) than for the negative clue (0.470).

c) Mean Scores of Same-Trials in the Nonreversal group

In order to determine if factors other than the reversal contributed to the score drops, the same analysis was also applied to the nonreversal group scores. The following table presents the summary statistics for same-trials of the nonreversal group of Experiment 3:

Table 17 Means, Standard Deviations and N for Accuracy Scores
of Nonreversal group for Same-Trials after 90

Nonreversal group: Same-Trials				
Clue	Trial Block	Accuracy Score (Mean)	Accuracy Score (Std. Dev.)	N ⁵
Positive	46-90	0.7512	0.2100	16
	91-120	0.6860	0.1783	13
	121-150	0.7642	0.1407	13
Negative	46-90	0.6354	0.3599	16
	91-120	0.5959	0.3298	13
	121-150	0.6087	0.3495	13
Neutral	46-90	0.6622	0.1596	16
	91-120	0.6474	0.1980	13
	121-150	0.6691	0.1809	13

⁵ The N values do not always match; because the clue type and the trial type was randomly determined for every trial of every session, trial type-clue type combinations were not equally distributed over trial blocks.

The following figure shows the changes in the mean scores listed in the above table:

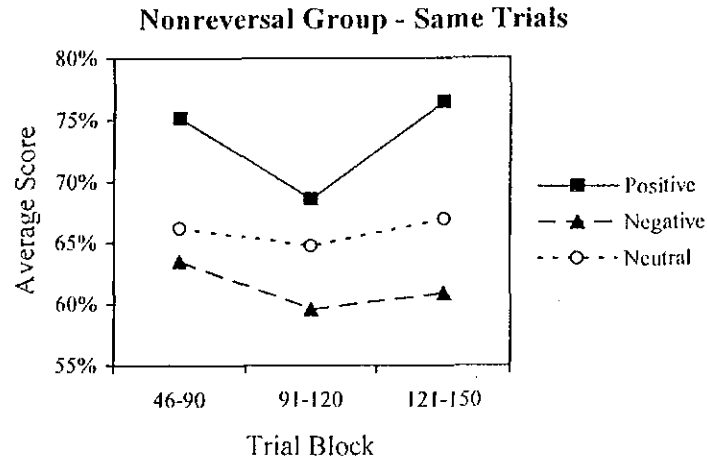


Figure 26 Mean Accuracy Scores of Nonreversal group for Same Trials after 90

The repeated-measures ANOVA test of same-trials in nonreversal group after 90 did not show a significant main effect of clue type ($F(2,24) = 2.096$, $MSe = 4.688$, $p = 0.145$, $r = 0.38$). However, when neutral clue was eliminated from the analyses a significant main effect of clue type was found ($F(1,12) = 3.792$, $MSe = 5.168$, $p = 0.075$, $r = 0.48$). The main effect of trial block was not significant. ($F(1,12) = 0.555$, $MSe = 4.966$, $p = 0.471$, $r = 0.20$). The interaction between clue type and trial block was not significant ($F(2,24) = 0.159$, $MSe = 5.149$, $p = 0.854$, $r = 0.11$). An evaluation of these results show that accuracy score of negative clue did not change much (mean = 0.5959 for trial block 91-120 and mean = 0.6087 for trial block 121-150). Accuracy score of the positive clue increased as expected (mean = 0.6860 for trial blocks 91-120 and mean = 0.7642 for trial blocks 121-150). When neutral clue trials were eliminated from the analyses, interaction was still not

significant ($F(1,12) = 0.225$, $p = 0.644$, $MSe = 6.188$, $r = 0.13$). There was no significant main effect of trial block $F(1,12) = 0.340$, $p = 0.570$, $MSe = 7.919$, $r = 0.16$. Overall, accuracy score was lower (0.641) for trial block 90-120 than for the trial block 121-150 (0.686). There was no significant main effect of clue type $F(1,12) = 3.792$, $p = 0.075$, $MSe = 5.168$, $r = 0.48$. Overall, accuracy score was higher for positive clue (0.725) than for the negative clue (0.602).

d) Mean Scores of Different-Trials in the Nonreversal group

The following table presents the summary statistics for different-trials of the nonreversal group:

Table 18 Means, Standard Deviations and N for Accuracy Scores
of Nonreversal group for Different-Trials after 90

Nonreversal group: Different-Trials				
Clue	Trial Block	Accuracy Score (Mean)	Accuracy Score (Std. Dev.)	N ⁶
Positive	46-90	0.5010	0.2210	15
	91-120	0.4693	0.2745	16
	121-150	0.5024	0.2239	16
Negative	46-90	0.5131	0.1596	15
	91-120	0.4467	0.2389	16
	121-150	0.5295	0.1765	16
Neutral	46-90	0.5011	0.3439	15
	91-120	0.4694	0.2879	16
	121-150	0.5156	0.3970	16

⁶ The N values do not always match; because the clue type and the trial type was randomly determined for every trial of every session, trial type-clue type combinations were not equally distributed over trial blocks.

The following figure shows the changes in the mean score listed in the above table:

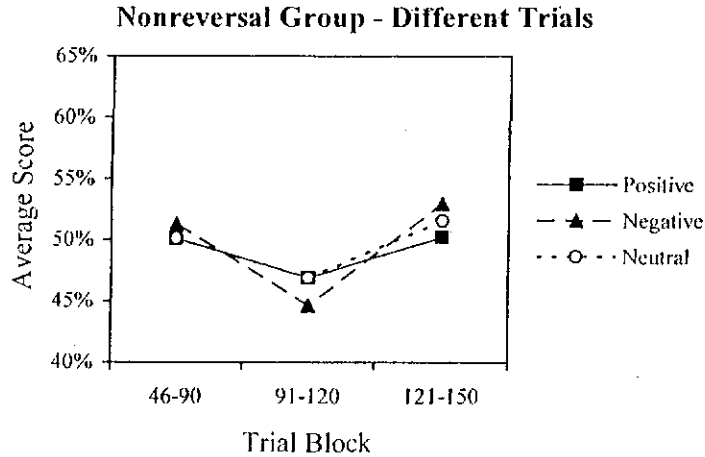


Figure 27 Mean Accuracy Scores of Nonreversal group for Different-Trials after 90

The repeated-measures ANOVA test of different-trials in the nonreversal group after 90 did not show a significant main effect of clue type ($F(2,30) = 0.005$, $MSe = 7.887$, $p = 0.995$, $r = 0$). The main effect of trial block was not significant either ($F(1,15) = 0.939$, $MSe = 7.468$, $p = 0.348$, $r = 0.24$). The interaction between clue type and trial block was not significant ($F(2,30) = 0.090$, $MSe = 5.913$, $p = 0.914$, $r = 0.07$). An evaluation of these results show that accuracy score of negative clue continued to increase as expected (mean = 0.4467 for trial block 91-120 and mean = 0.5295 for trial block 121-150). However, the accuracy score of the positive clue also increased (mean = 0.4693 for trial block 91-120 and mean = 0.5024 for trial block 121-150). When neutral clue trials were eliminated from the analyses, interaction was still not significant ($F(1,15) = 0.199$, $p = 0.662$, $MSe = 4.965$, $r = 0.11$). There was no significant main effect of trial block $F(1,15) = 0.882$, $p = 0.362$,

MSe= 6.090, $r = 0.32$). Overall, accuracy score was little lower (0.458) for trial block 90-120 than for the trail block 121-150 (0.516). There was no significant main effect of clue type $F(1,15) = 0.002$, $p = 0.968$, MSe= 4.842, $r = 0$). Overall, accuracy score was little lower for positive clue (0.486) than for the negative clue (0.488).

C. Summary of the Results

The changes in the task performance of the participants of experiment 3 until the 90th trial show evidence of clue learning. Mean accuracy scores were high, about 70%, for same-trials, but lower for different-trials. In same trials, the mean accuracy score for positive clue trials increased as expected (mean = 0.7378 for trial block 1-45 and mean = 0.7711 for trial block 46-90). The average score for negative clue trials decreased (mean = 0.8060 for trial block 1-45 and mean = 0.6693 for trial block 46-90) also as expected. As for the different trials, mean score for positive clue was 0.5141 for the first trial block (1-45) and 0.5005 for the second trial block (46-90). However, the mean scores of those blocks differed little. In addition, mean score for negative clue was 0.5158 for the first trial block and 0.5388 for the second trial block, an increase as expected, but an insignificant one.

When we compare the mean scores of reversal group and nonreversal group we also see the evidence of implicit learning of the rules. Nonreversal group scores after 90th trial continued to fulfill our expectations, especially same trials. The results showed that the mean accuracy score of trials with negative clues did not change in same trials. The mean

score for trials with positive clue continued to increase as expected in same-trials. In other words, accuracy scores of the same and different-trials continued to stay different from each other throughout the whole session (150 trials). As for the different trials, the mean accuracy score of negative clue also continued to increase as expected. However, the mean accuracy score of the positive clue also increased. The reversal group scores after the 90th trial does not show an expected score drop however, the lack of significant score drop in the reversal group can be attributed to a quick re-learning of the new predictiveness relationship in different trials.

The participants' responses to the post-task questionnaires indicated that none of the participants were performing the task by explicitly relying on the predictiveness relationships of the clues. None of the participants reported having guessed or discovered those relationships.

Chapter 7

THE METHODOLOGY AND RESULTS OF EXPERIMENT 4

XIII. OVERVIEW OF EXPERIMENT 4

A. Purpose

The purpose of Experiment 4 was to examine the process of implicit learning in auditory perception by utilizing the predictive context of rhythm clues embedded in musical melodies.

B. Method

The main task in Experiment 4 was the same task commonly used throughout the present study: determining whether or not a pair of melodies was same or different. Rhythm clues, i.e., sequences of notes with unchanging lengths were embedded in melodies provided predictive information relating to whether or not a specific pair was likely to be a same-pair or a different-pair. Three different rhythm clues (see page 35) of three different predictive relationships with same-trials (75% for positive clue, 50% for neutral clue and 25% for negative clue) were used in equal proportions throughout 150 trials that made up each experimental session. In Experiment 4, melodies were played at a fast tempo (200 beats per minute, a standard note lasts 1/200 of a minute) to make melodies

sound more like complete patterns. Each melody (including its clue) lasted about 4 seconds.

C. Participants

A total of 32 people participated in this experiment, 13 in the reversal group, and 19 in the nonreversal group. The numbers of participants exceeding the score limit were 10 in the reversal group and 19 in the nonreversal group.

Each participant who completed an experimental session received two hours credit for a course requirement.

XIV. RESULTS OF EXPERIMENT 4

A. Analysis of Performance Changes before the 90th Trial

1. Mean Scores of Same-Trials

The following table shows the summary statistics for the accuracy scores of Experiment 4 for same-trials:

Table 19 Means, Standard Deviations and N for Accuracy Scores for Same-Trials before 90

Same-Trials				
Clue	Trial Block	Accuracy Score (Mean)	Accuracy Score (Std. Dev.)	N
Positive	1-45	0.7222	0.1846	29
	46-90	0.7379	0.1996	29
Negative	1-45	0.7391	0.3107	29
	46-90	0.8017	0.2533	29
Neutral	1-45	0.7596	0.1876	29
	46-90	0.7531	0.1546	29

The following figure shows the changes in the mean accuracy scores summarized in the above table:

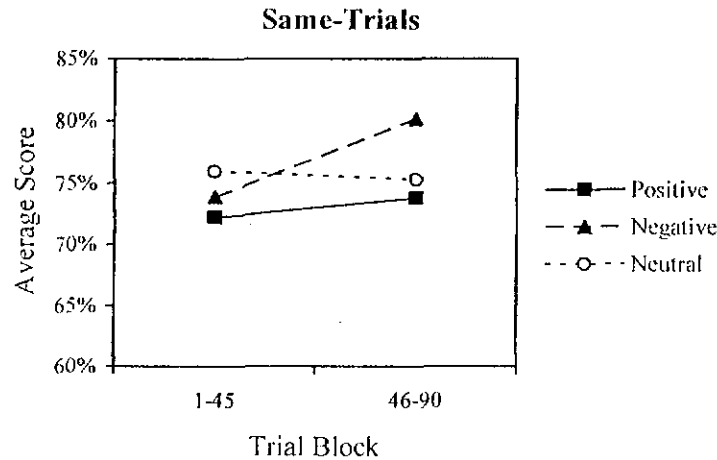


Figure 28 Mean Accuracy Scores of Same-Trials before 90

According to 3 (clue type) by 2 (trial block) repeated-measures ANOVA test, the expected performance improvements are not observed in same-trials. There is no significant main effect of clue type ($F(2,56) = 0.080$, $MSe=5.281$, $p = 0.924$, $r = 0.05$). No significant main effect is found for trial block either ($F(1,28) = 0.066$, $MSe = 3.494$, $p = 0.799$, $r = 0.04$). No significant interaction was found between the clue type and the trial block ($F(2,56) = 1.296$, $MSe = 2.786$, $p = 0.282$, $r = 0.20$). When neutral clue trials were eliminated from the analyses, the interaction was still not significant ($F(1,28) = 2.123$, $MSe = 3.206$, $p = 0.156$, $r = 0.26$). There was no significant main effect for clue type. ($F(1,28) = 0.096$, $MSe= 7.099$, $p = 0.758$, $r = 0.05$). Overall participants did better with negative clue (0.770) than positive clue (0.755). There was no significant main effect for trial block ($F(1,28) = 0.123$, $MSe = 4.741$, $p = 0.728$, $r = 0.06$). Overall, participants in block 1-45 did

worst than (marginal mean= 0.756) and the participants in block 45-90 (marginal mean= 0.770). Out of 29 participants, 9 participants' accuracy scores increased for positive clue. As for the negative clue, 16 participants' accuracy scores increased.

An evaluation of the cell means showed that, the mean accuracy score for positive clue does not change much (mean = 0.7222 for trial block 1-45 and mean = 0.7379 for trial block 46-90). In addition, the average score for negative clue increases instead of decreasing (mean = 0.7391 for trial block 1-45 and mean = 0.8017 for trial block 46-90).

2. Mean Scores of Different-Trials

The summary statistics for the accuracy scores for different-trials are shown in the following table:

Table 20 Means, Standard Deviations and N for Accuracy Scores for Different-Trials before 90

Different-Trials				
Clue	Trial Block	Accuracy Score (Mean)	Accuracy Score (Std. Dev.)	N
Positive	1-45	0.6264	0.2202	29
	46-90	0.5627	0.2387	29
Negative	1-45	0.5801	0.2002	29
	46-90	0.6209	0.1384	29
Neutral	1-45	0.4736	0.2766	29
	46-90	0.4598	0.3144	29

The following figure shows the changes in the mean scores listed in the above table:

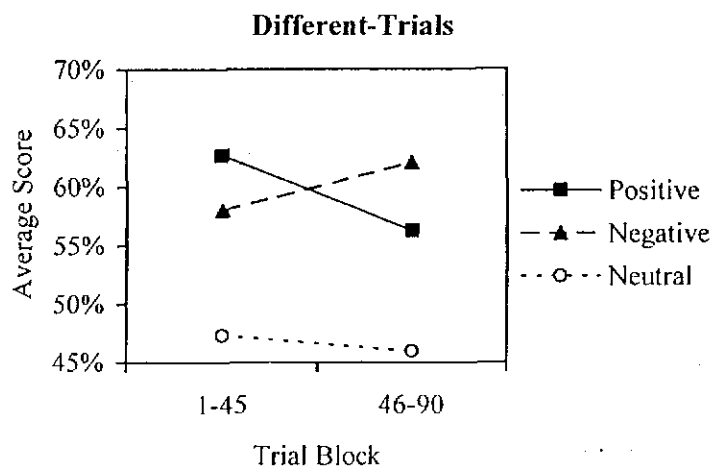


Figure 29 Mean Accuracy Scores of Different-Trials before 90

According to 3 (clue type) by 2 (trial block) repeated-measures ANOVA test, the expected performance improvements were observed for different-trials. Mean score for the positive clue was 0.6264 for the first trial block (1-45) and 0.5627 for the second trial block (46-90), a decrease as expected. Mean score for negative clue was 0.5801 for the first trial block and 0.6209 for the second trial block, an increase as expected, but an insignificant one. According to the ANOVA test, there is a significant main effect of clue type ($F(2,56) = 9.901$, $MSe = 3.348$, $p = 0.000$, $r = 0.51$), however this effect is due to the inclusion of the neutral clue. There was no significant main effect of trial block ($F(1,28) = 0.150$, $MSe = 4.306$, $p = 0.701$, $r = 0.07$). Finally, there was no significant interaction to be found between clue type and trial block ($F(2,56) = 0.960$, $MSe = 4.127$, $p = 0.389$, $r = 0.18$). When neutral clue trials were eliminated from the analyses, interaction still was not significant. ($F(1,28) = 2.017$, $MSe = 3.926$, $p = 0.167$, $r = 0.25$). There was no significant main effect for clue type ($F(1,28) = 0.041$, $MSe = 2.500$, $p = 0.840$, $r = 0.03$). However, as hypothesized, overall participants did better with the negative clue (0.601) than did the

positive clue (0.595). There was no significant main effect for trial block ($F(1,28) = 0.199$, $MSe = 1.899$, $p = 0.659$, $r = 0.08$). However, this lack of a main effect is the result offsetting simple effects of trial block for the positive clue which showed an decrease in average score from block1 to block2 as expected and of trial block for the negative clue which showed an increase in average score from block1 to block2. Overall, participants in block 1-45 (marginal mean= 0.603) did better than the participants in block 45-90 (marginal mean= 0.592). Out of 29 participants, 15 participants' accuracy scores increased for negative clue. As for the positive clue, only 12 participants' accuracy scores increased.

3. Mean Scores of Positive-Clue Trials

The comparison of accuracy scores of same-trials and different-trials with positive clue showed that the changes in them were in the expected directions.

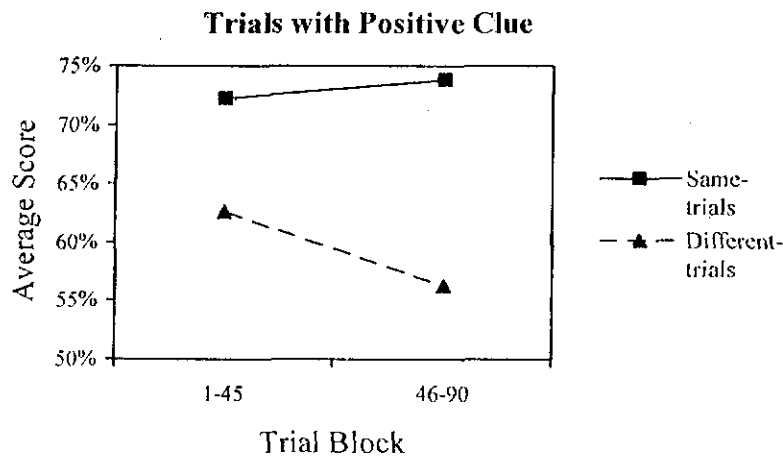


Figure 30 Mean Accuracy Scores of Positive Clue for Trials before 90

The repeated-measures ANOVA test shows no significant main effect of trial block for trials starting with the positive clue ($F(1,28) = 2.243$, $MSe = 3.098$, $p = 0.145$, $r = 0.27$). But there is a significant main effect of trial type (i.e., same or different) ($F(1,28) = 11.238$, $MSe = 6.648$, $p = 0.002$, $r = 0.53$). As hypothesized, accuracy scores were higher for same trials (mean= 0.755) than for the different trials (mean= 0.595). No significant interaction between trial type and trial block ($F(1,28) = 0.173$, $MSe = 3.618$, $p = 0.680$, $r = 0.07$).

4. Mean Scores of Negative-Clue Trials

When the accuracy scores of same-trials and different-trials of negative clue were compared, the following figure was obtained: As seen in Figure 35, negative clue increases in same trials instead of decreasing.

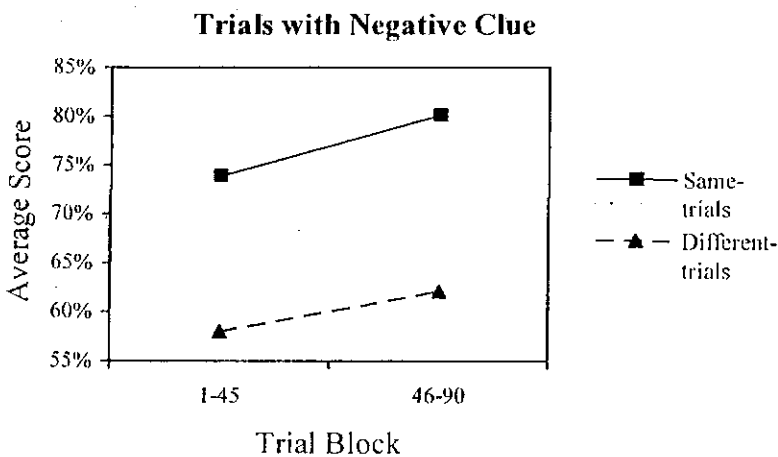


Figure 31 Mean Accuracy Scores of Negative Clue for Trials before 90

The repeated-measures ANOVA test shows no significant main effect of trial block for trials starting with the negative clue ($F(1,28) = 2.064$, $MSe = 3.763$, $p = 0.162$, $r = 0.26$). But there is a significant main effect of trial type (i.e., same or different) ($F(1,28) = 9.858$, $MSe = 8.492$, $p = 0.004$, $r = 0.50$). Average accuracy score is higher for same trials (mean= 0.770) than for different trials (mean= 0.601) which is the opposite of the hypothesized pattern. No significant interaction between trial type and trial block ($F(1,28) = 0.105$, $MSe = 3.292$, $p = 0.749$, $r = 0.06$).

B. Analysis of Performance Changes after the 90th Trial

1. Score Drops In the Experiment and Nonreversal groups

In order to determine if there was a performance cost associated with the reversal of the predictiveness relationships, it was necessary to calculate the mean accuracy scores separately for the reversal-group and nonreversal group. The performance cost of the reversal is apparent in the substantial drops in the mean scores for all clues in both participant groups. In order to determine the significance of the score drops observed for positive and negative clues, the mean scores of those clues were analyzed together for the trial blocks 46-90 and 91-120. The following table lists the combined mean scores of same-trials and different-trials for those two blocks:

Table 21 Means, Standard Deviations and N for Accuracy Scores
of Experiment and Nonreversal groups for Trials before and after 90

Participant Group	Trial Blocks	
	46-90	91-120
Reversal group	Mean = 0.6874	Mean = 0.6597
	Std. Dev. = 0.2169	Std. Dev. = 0.2672
	N = 40	N = 40
Nonreversal group	Mean = 0.6757	Mean = 0.6858
	Std. Dev. = 0.2388	Std. Dev. = 0.2327
	N = 75	N = 75

There is a decrease in accuracy scores of reversal group after the reversal. However, the changes are found not to be significant for both the reversal group ($t(39) = 0.604$, $p > 0.05$) and for the nonreversal group ($t(74) = -0.333$, $p > 0.05$).

2. Mean Scores of Trials after 90

a) Mean Scores of Same-Trials in the Reversal group

The following table summarizes the summary statistics for same-trials of reversal group:

Table 22 Means, Standard Deviations and N for Accuracy Scores
of Reversal group for Same-Trials after 90

Reversal group: Same-Trials				
Clue	Trial Block	Accuracy Score (Mean)	Accuracy Score (Std. Dev.)	N
Positive	46-90	0.6804	0.2454	10
	91-120	0.6767	0.2179	10
	121-150	0.8067	0.1953	10
Negative	46-90	0.7700	0.2574	10
	91-120	0.8156	0.1923	10
	121-150	0.7105	0.2035	10
Neutral	46-90	0.6909	0.1636	10
	91-120	0.6583	0.3202	10
	121-150	0.7167	0.3429	10

The following figure shows the changes in the mean accuracy scores summarized in the above table:

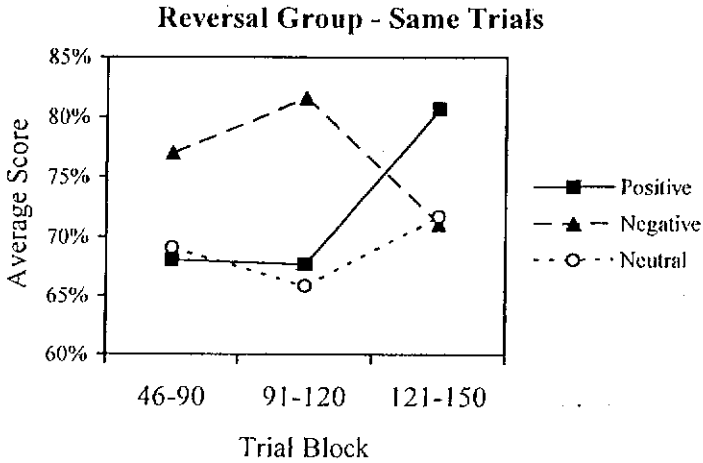


Figure 32 Mean Accuracy Scores of Reversal group for Same-Trials after 90

The repeated-measures ANOVA test of same-trials in reversal group after 90 did not show a significant main effect of clue type ($F(2,18) = 0.568$, $MSe = 5.341$, $p = 0.576$, $r = 0.24$). The main effect of trial block was not significant either ($F(1,9) = 0.116$, $MSe = 9.934$, $p = 0.741$, $r = 0.11$). The interaction between clue type and trial block was not significant ($F(2,18) = 1.427$, $MSe = 5.088$, $p = 0.266$, $r = 0.37$). When neutral clue trials were excluded from the analyses, interaction became significant ($F(1,19) = 6.709$, $MSe = 3.438$, $p = 0.019$, $r = 0.53$). However, inspection of cell means showed that accuracy score of negative clue decreased (mean = 0.8156 for trial block 91-120 and mean = 0.7105 for trial block 121-150), instead of increasing. Accuracy score of the positive clue increased (mean = 0.6767 for trial block 91-120 and mean = 0.8067 for trial block 121-150), instead of decreasing.

b) Mean Scores of Different-Trials in the Reversal group

The following table summarizes the summary statistics for different-trials of the reversal group:

Table 23 Means, Standard Deviations and N for Accuracy Scores
of Reversal group for Different-Trials after 90

Reversal group: Different-Trials				
Clue	Trial Block	Accuracy Score (Mean)	Accuracy Score (Std. Dev.)	N
Positive	46-90	0.6698	0.1735	10
	91-120	0.5217	0.1934	10
	121-150	0.5533	0.2744	10
Negative	45-90	0.6295	0.1882	10
	91-120	0.6250	0.3711	10
	121-150	0.5667	0.3187	10
Neutral	45-90	0.6017	0.3210	10
	91-120	0.4965	0.2285	10
	121-150	0.5092	0.1518	10

The following figure shows the changes in the mean accuracy scores summarized in the above table:

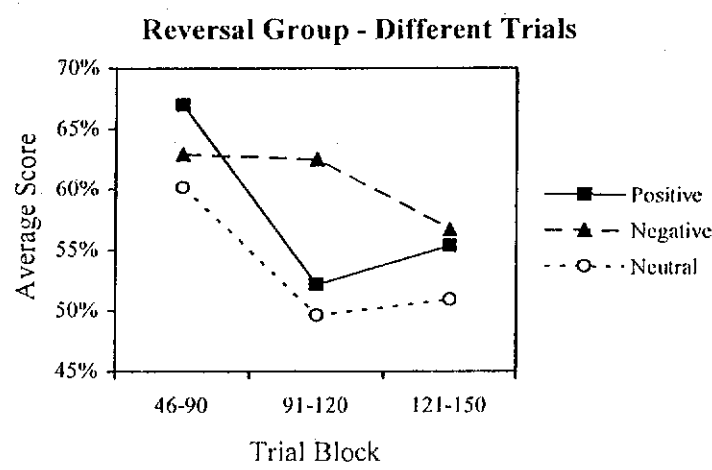


Figure 33 Mean Accuracy Scores of Reversal group for Different-Trials after 90

The repeated-measures ANOVA test of different-trials in the reversal group after 90 did not show a significant main effect of clue type ($F(2,18) = 0.963$, $MSe = 4.589$, $p > 0.05$, $r = 0.31$). The main effect of trial block was not significant either ($F(1,9) = 0.006$, $MSe = 5.082$, $p > 0.05$, $r = 0.03$). The interaction between clue type and trial block was not significant ($F(2,18) = 0.255$, $MSe = 4.418$, $p > 0.05$, $r = 0.16$). When neutral clue trials were eliminated from the analyses, interaction was still not significant ($F(1,9) = 0.257$, $p = 0.619$, $MSe = 2.153$, $r = 0.11$). There was no significant main effect of trial block $F(1,9) = 0.839$, $p = 0.372$, $MSe = 4.129$, $r = 0.21$). Overall, accuracy score was lower (0.751) for trial block 90-120 than for the trail block 121-150 (0.819). There was no significant main effect of clue type $F(1,9) = 0.364$, $p = 0.554$, $MSe = 3.868$, $r = 0.14$). Overall, accuracy score was higher for positive clue (0.798) than for the negative clue (0.773).

These results show that accuracy scores of negative clue trials continued to decrease as expected (mean = 0.6250 for trial block 91-120 and mean = 0.5667 for trial block 121-150). The mean accuracy scores of the positive clue trials also continued to increase as expected (mean = 0.5217 for trial block 91-120 and mean = 0.5533 for trial block 121-150).

c) Mean Scores of Same-Trials in the Nonreversal group

The following table presents the summary statistics for same-trials of the nonreversal group of Experiment 4:

Table 24 Means, Standard Deviations and N for Accuracy Scores
of Nonreversal group for Same-Trials after 90

Nonreversal group: Same-Trials				
Clue	Trial Block	Accuracy Score (Mean)	Accuracy Score (Std. Dev.)	N ⁷
Positive	46-90	0.7682	0.1704	19
	91-120	0.7504	0.2311	18
	121-150	0.8693	0.1198	18
Negative	46-90	0.8184	0.2565	19
	91-120	0.7477	0.2073	18
	121-150	0.6402	0.3786	18
Neutral	46-90	0.7858	0.1431	19
	91-120	0.8223	0.1992	18
	121-150	0.7937	0.2444	18

⁷ The N values do not always match; because the clue type and the trial type was randomly determined for every trial of every session, trial type-clue type combinations were not equally distributed over trial blocks.

The following figure shows the changes in the mean scores listed in the above table:

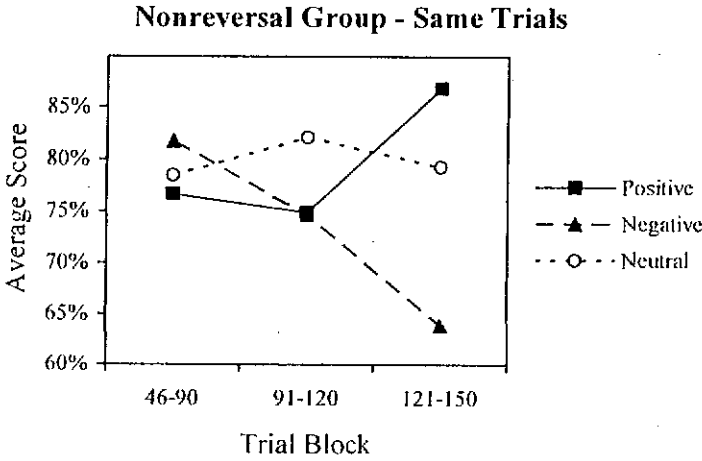


Figure 34 Mean Accuracy Scores of Nonreversal group for Same Trials after 90

The repeated-measures ANOVA test of same-trials in nonreversal group after 90 did show a significant main effect of clue type ($F(2,34) = 3.369$, $MSe = 4.709$, $p = 0.046$, $r = 0.40$). The main effect of trial block was not significant. ($F(1,17) = 0.14$, $MSe = 6.573$, $p = 0.909$, $r = 0.03$). The interaction between clue type and trial block was not significant ($F(2,34) = 2.573$, $MSe = 4.619$, $p = 0.091$, $r = 0.36$). However when neutral clue trials were excluded from the analyses, the interaction became significant ($F(1,17) = 6.709$, $MSe=3.438$, $p = 0.019$, $r = 0.53$). There was no significant main effect of trial block ($F(1, 17) = 0.08$, $MSe= 7.202$, $p = 0.929$, $r = 0$). There was a significant main effect of clue type ($F(1,17)= 4.788$, $MSe= 5.051$, $p = 0.43$, $r = 0.20$). As hypothesized, overall, accuracy score was higher for the positive clue (0.810) than for the negative clue (0.694). Inspection of the cell means also showed that accuracy score of negative clue continued to decrease as expected (mean = 0.7477 for trial block 91-120 and mean = 0.6402 for trial block 121-

150). Accuracy score of the positive clue also continued to increase as expected (mean = 0.7502 for trial block 91-120 and mean = 0.8693 for trial block 121-150).

d) Mean Scores of Different-Trials in the Nonreversal group

The following table presents the summary statistics for different-trials of the nonreversal group:

Table 25 Means, Standard Deviations and N for Accuracy Scores
of Nonreversal group for Different-Trials after 90

Nonreversal group: Different-Trials				
Clue	Trial Block	Accuracy Score (Mean)	Accuracy Score (Std. Dev.)	N ⁸
Positive	46-90	0.5063	0.2526	19
	91-120	0.6427	0.2326	18
	121-150	0.6271	0.2907	18
Negative	46-90	0.6164	0.1097	19
	91-120	0.5872	0.2136	18
	121-150	0.5337	0.2972	18
Neutral	46-90	0.4600	0.2918	19
	91-120	0.4944	0.3402	18
	121-150	0.5431	0.3451	18

⁸ The N values do not always match; because the clue type and the trial type was randomly determined for every trial of every session, trial type-clue type combinations were not equally distributed over trial blocks.

The following figure shows the changes in the mean score listed in the above table:

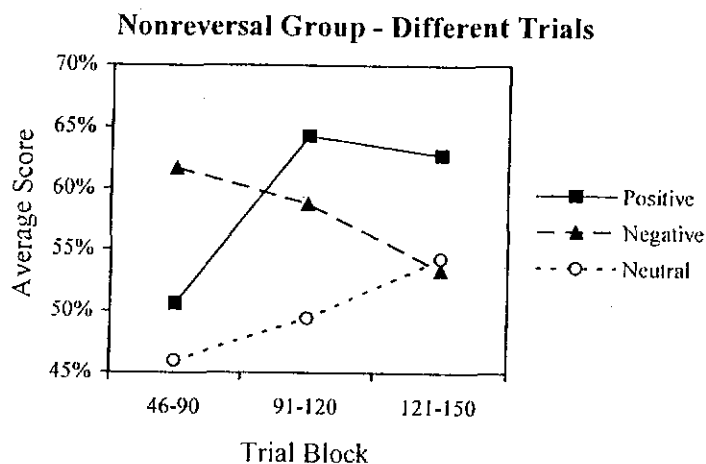


Figure 35 Mean Accuracy Scores of Nonreversal group for Different-Trials after 90

The repeated-measures ANOVA test of different-trials in the nonreversal group after 90 did not show a significant main effect of clue type ($F(2,34) = 1.344$, $MSe = 9.273$, $p = 0.274$, $r = 0.27$). The main effect of trial block was not significant either ($F(1,17) = 0.022$, $MSe = 5.619$, $p = 0.883$, $r = 0.1$). The interaction between clue type and trial block was not significant ($F(2,34) = 0.398$, $MSe = 6.036$, $p = 0.675$, $r = 0.15$). When neutral clue trials were excluded from the analyses, the interaction was still not significant ($F(1,17) = 0.257$, $MSe = 4.455$, $p = 0.619$, $r = 0.11$). There was no significant main effect of trial block ($F(1, 17) = 0.839$, $MSe = 4.324$, $p = 0.372$, $r = 0.21$). There was no significant main effect of clue type ($F(1,17) = 0.364$, $MSe = 9.243$, $p = 0.554$, $r = 0.14$). Accuracy score was higher for positive clue (0.617) overall than for the negative clue (0.575). Inspection of the cell means also showed that accuracy score of negative clue continued to decrease (mean =

0.5872 for trial block 91-120 and mean=0.5337 for trial block 121-150).

Accuracy score of the positive clue increased a little (mean = 0.6427 for trial block 91-120 and mean = 0.6271 for trial block 121-150).

C. Summary of the Results

The analyses of the task performance measures before the 90th trial showed only weak evidence of clue learning. In some trials, the mean scores for positive clue did not change much. In addition, the average score for negative clue increased instead of decreasing. In other words, there is not much to say about any clue learning at all. However, in different trials, the expected performance improvements were observed for both positive clue and negative clue.

The reversal of predictiveness relationships did not lead to the expected performance cost among the reversal group participants. There was more score drop among nonreversal group participants, meaning that factors other than the reversal were more effective in the task performance. For reversal group participants, only the mean score of different trials with positive clue decreased as expected; the mean scores remained the same for the negative clue. Reversal group participants also demonstrated that they re-learned the clues' roles after the reversal; their mean scores changed in the expected direction after the 90th trial. The mean accuracy scores show that accuracy score of negative clue trials continued to decrease as expected. The mean accuracy score of the positive clue trials also continued to increase as expected in trial blocks 91-120 and 121-150. However, we do not see the expected pattern of learning in the scores of non-reversal group. In other

words, different trials do not exhibit the clue learning in nonreversal group but only in reversal group. Same trials do not show clue learning for either group.

The participants' responses to the post-task questionnaires conclusively eliminated the possibility that any of the participants were performing the task by explicitly relying on the predictiveness relationships of the clues. None of the participants reported having guessed or discovered those relationships.

Chapter 8

DISCUSSIONS AND CONCLUSIONS

XV. DISCUSSIONS ON THE RESULTS OF THE EXPERIMENTS

A. A Discussion on the Results of Experiment 1

Neither the reversal-group participants nor the nonreversal-group participants of Experiment 1 showed the expected patterns that would result from the implicit learning of the roles of clues. The changes observed in the mean scores of reversal group verified that the reversal of predictiveness relationships had a net performance cost, but not having observed the expected score changes before the reversal, the performance cost by itself was not considered a sufficient evidence for the presence of implicit learning.

The lack of sufficient evidence may be because the melody clues were too explicit when the melody pairs were played in slow tempo. The apparent presence of repeating patterns at the beginning of every melody might have led many participants to initially attribute some meaning to clues (indeed, responses of some participants to post-task questionnaires mentioned that clues may have had meanings), but not the right meanings (no participant's responses indicated that he had found the correct relationship). Having not discovered the correct relationships, participants would have had to detect the different notes in different trials to succeed, and judging from the generally low accuracy scores, this

was a difficult feat to perform. The main reason for the difficulty may be the loss of attention due to the long duration of each trial and the whole session.

B. A Discussion on the Results of Experiment 2

The changes in the task performance of the participants of Experiment 2 until the 90th trial were in the right directions expected from implicit learning of the predictive context. Although the changes were not found to be significant with respect to trial, they were significantly different for different clues. When trials with positive clue were analyzed separately from the others, changes in the mean scores of same-trials and negative-trials were found to be in sharp contrast. A similar contrast was found between the scores of same-trials and different-trials starting with the negative clue.

The drop in the mean accuracy score after the 90th trial was greater in the nonreversal group than in the reversal group, but the lack of significant drops in the reversal groups are attributed to a quick re-learning of the reversed predictiveness level of positive and negative clues in the same trials.

Melody clues were seemingly less explicit when melody pairs were played with a fast tempo. Thus the participants of Experiment 2 probably concentrated on performing the task without attributing arbitrary meanings to clues. Nevertheless, there is sufficient evidence that implicit learning occurred, and Experiment-2 participants did perform better than Experiment-1 participants. As the background review of implicit learning studies made it clear, non-obviousness of the rules underlying the task structure (in the present study, the predictive relationships) was indeed one of the prerequisites of implicit learning.

Experiment 2 showed that if specific note patterns are to be used to promote implicit learning through auditory stimuli, the playback tempo must be carefully adjusted to prevent the note patterns from being too apparent.

C. A Discussion on the Results of Experiment 3

The results presented so far indicate that participants of Experiment 3 implicitly learned the predictiveness relationships of clues. The only contradictory result was that the reversal-group participants did not experience a significant score drop as a result of the reversal of the predictiveness relationships. A strong increase in the mean score of trials with positive clue and a strong decrease in the mean score of trials with negative clue in different trials implied that reversal-group participants re-learned the new, reversed role of clues faster than expected. The re-learning effect was probably the reason why the score drops in the reversal group was weaker than expected.

Significant score drops observed in the nonreversal group indicated that certain unexpected factors affected the performances of all participants. The difficulty of the task and the growing lack of attention may have been among those factors. Thus the result of Experiment 3 points to the need for a careful re-design of the task in future experiments intended to study implicit learning utilizing musical melodies.

Interestingly, Experiment-3 participants performed better than Experiment-1 participants, even though melodies were played with a slow tempo in both experiments. Thus the results from Experiment 3 imply that rhythm clues were also effective tools to promote implicit learning, even if the playback tempo was slow, because rhythm clues did

not stand out as much as melody clues. This is again another observation supporting the conclusion that auditory clues in experiments involving music perception must not be made explicit and melodies must sound like complete patterns.

D. A Discussion on the Results of Experiment 4

The analyses of task performance measures before the 90th trial showed only weak evidences indicating the presence of implicit learning. The mean scores changed in the expected ways for the trials with positive clue, but those changes were quite weak. The mean score of trials with negative clue did not change according to the predicted patterns. Finally, the reversal of the predictiveness relationships did not lead to the expected performance cost among the reversal-group participants and the score drops were more significant than expected among nonreversal-group participants.

When they are contrasted with the results of Experiment 3, the results of Experiment 4 indicate that rhythm clues did not help implicit learning develop when the melody stimuli were played at a fast tempo. It should be noted that rhythm clues, when played at a fast tempo like they were in Experiment 4 stimuli, were completely imperceptible; fast-tempo melodies sounded like complete, unbroken patterns with their rhythm clues. This is a conclusion opposite to the drawn from the results of Experiment 2, which implied that the melody clues helped implicit learning when the auditory patterns were complete. In other words, the occurrence of implicit learning in an experimental task involving melodies depends on how well the auditory clues fit within the melodies. If the clues are too apparent, they will promote explicit learning rather than implicit learning.

That may have been the reason why no evidence of implicit learning was found in Experiment 1. On the other hand, when the clues fit too well within the melodies like in Experiment 4, again no indications of implicit learning were found, because participants never perceived the clues as possible carriers of useful information.

E. A Discussion on the Responses to Questionnaires

The participants' responses to the post-task questionnaires conclusively eliminated the possibility that any of the participants were performing the task by explicitly relying on the predictiveness relationships of the clues. The responses of a few participants indicated that they were aware of repeating patterns at the beginnings of the melodies, if melody clues were present in them. However, even those participants did not acknowledge that they attributed any meanings to those patterns or indicated that they discovered the patterns served as predictive clues.

XVI. CONCLUSIONS

Certain characteristics of the four experiments described in this paper made them especially unique in the field of implicit learning. Firstly, the experiments making up this study employed an auditory perception task utilizing the probabilistic relationships between the structural elements of music. That was a distinguishing characteristic because very few or no implicit learning studies dealt with music directly. Secondly, the primary experimental construct targeted by data analysis was the accuracy score, whereas most implicit learning studies concentrated on analyzing the reaction times, because accuracy scores observed in those experiments were either too high or showed too little variation. In this study, however, reaction time showed too little variation to be useful as a performance measure and accuracy score showed enough variation to be a better measure of performance.

According to the results discussed so far, participants of Experiments 2 and 3 learned the predictiveness relationships of clues embedded in the melodies. Since responses to questionnaires eliminated the possibility of explicit learning, it was concluded that the participants of those two experiments had learned the roles of the clues implicitly. Therefore, the types of relationships utilized in Experiments 2 and 3 were concluded to be promoting implicit learning.

Experiment 2 employed specific note patterns (i.e., melody clues) as clues in melodies played at a fast tempo. However, Experiment 3 employed rhythm clues consisting of different notes for almost every trial and melodies in that experiment were played at a

slow tempo. It looks like sensitivity to context was playing a big role in the occurrence of implicit learning in auditory perception. In a study done by Flowers and Grafel (2002), it was found that the auditory stimuli used to convey the information first presented in visual displays were found to be more effective when they are played faster (7.1 seconds) than slower (14.2 seconds). Since the results of this study also showed that the exact durations of melody stimuli affected the occurrence of implicit learning as well as the types of the clues embedded in melodies, the observations of Flowers and Grafel (2002) are especially relevant for this study.

It is clear that, not only the type of structural elements used in musical stimuli, but also the durations of those stimuli affect the occurrence of learning, whether it is implicit learning of relationships or learning of visual displays transformed into musical melodies. For example, if the durations of melodies are too long as in Experiment 1, the experimental task may become too difficult to perform and learning will be nearly impossible. However, when the same melodies were presented faster and thus shorter forms in Experiment 2, participants were able to learn the role of clues implicitly, indicating that the occurrence of implicit learning was negatively affected by increasing duration of melodies, at least over a certain range.

On the other hand, when rhythm clues were employed, slower melodies were found to be more effective than faster melodies as the results of Experiments 3 and 4 showed. Although both experiments used rhythm clues embedded in the same melodies, no significant evidence of implicit learning was observed in Experiment 4, whereas the results

of Experiment 3 suggested that rhythm clues did promote implicit learning when the playback tempo was slow.

The common point between the experiments 2 and 3 was that, in both experiments, melodies with clues sounded more like complete patterns in which clues were only slightly perceptible. The lack of evidence for implicit learning in Experiments 1 was attributed to the apparentness of the melody clues, whereas the absence of implicit learning in Experiment 4 may be due to the imperceptibility of the rhythm clues in shorter melodies. In other words, the durations of auditory stimuli must be chosen carefully so that people can keep a representative form of information presented via the stimuli in both sensory memory and in the working memory. The effects of structure of musical patterns in human mind must not be forgotten either. In short, this study shows that implicit learning of information or relationships conveyed through auditory stimuli may require a combination of consideration of both the temporal constraints of auditory sensory and working memory, and the perceptibility of the musical patterns.

Naturally, being one of the first studies in its field, this study also had significant shortcomings. For one thing, the memory task employed in the experiments proved to be too difficult to perform for many participants. This was apparent from the fact that the data from close to half of the participants were left out because the mean scores were not beyond the levels that could be achieved by chance alone. Another thing is that, the experimental sessions were probably too long. The score drops observed in non-reversal sessions were attributed to decreasing attention of participants, since those sessions did not involve any change in the predictiveness relationships of the clues. Thus the results of this

study have important implications relating to the directions that can be taken by future studies on the same subject. It is clear that implicit learning tasks involving musical melodies must be straightforward, interactive and short-lasting. The difficulty of an experimental task normally makes it more likely that participants of the experiment will rely on implicit learning while performing the task. However, as the results of this study showed, when music is used as the primary carrier of information, the difficulty of the task may in fact prove to be a detrimental factor preventing the occurrence of implicit learning. Therefore, future experiments employing melodies may have to require rapid and accurate responses from the participants to ensure that they will respond by learning the task structure implicitly. In such experiments, reaction time rather than accuracy score will probably be the primary indicator of implicit learning.

REFERENCES

- o, R. (1994) Music and language: Parallels and contrasts. In R. Aiello and J. A. Sloboda, (Eds), Musical Perceptions. Oxford University Press: New York.
- att, J.C. (1993) Tonal structure of melodies. In T. J. Tighe, & W. J. Dowling, (Eds.), Psychology and Music The Understanding of melody and Rythm. (pp.50). Lawrence Erlbaum Associates: New Jersey.
- , D. (1997). How implicit learning is implicit learning? Oxford University Press: Oxford.
- D. (1993). The Control of complex systems. In D. Berry, & Z. Dienes, (Eds.), Implicit Learning, Theoretcal and Empirical (pp.20) Lawrence Erlbaum Associates, Ltd.
- D., & Dienes, Z. (1993). Toward a working characterisation of implicit learning. In D. Berry, & Z. Dienes (Eds.), Implicit Learning, Theoretical and Empirical. (pp. 14). Lawrence Erlbaum Associates, Ltd.
- , D.C., & Broadbent, D.E. (1984). On the relationship between task performance and associated verbalizable knowledge. Quarterly Journal of Experimental Psychology, 36, 209-231.
- , D.C., & Broadbent, D.E. (1987). Explanation and verbalization in a computer-assisted search task. Quarterly Journal of Experimental Psychology, 39A6, 585-609.
- , D.C., & Broadbent, D.E. (1988). Interactive tasks and the implicit-explicit distinction. British Journal of Psyhology, 79, 251-272.
- lbent, D.E. (1977). Levels, hierarchies and the locus of control. Quarterly Journal of Experimental Psychology, 29, 181-201.

- deBent, D.E., & Aston, B. (1978). Human control of a simulated economic system. Ergonomics, 21, 1035-1043.
- Deutsch, E.M. (1999). Intervals, scales, and tuning. In D. Deutsch, (Ed.), The Psychology of music. (pp. 215- 216). Academic Press : San Diego.
- Deutsch, D. (1992). The Musician's guide to perception and cognition. Schirmer Books: New York.
- Deutsch, E.C., & Kendall, R.A. (1999) Comparative music perception and cognition. D. Deutsch, (Ed.), The Psychology of music. Academic Press: San Diego.
- Chomsky, N., & Miller, G.A. (1958). Finite State Languages. Information and Control, In A. S. Reber, (Eds.), Learning. about complex sets of rules.
- Remmans, A., & Jimenez, L. (1998) Implicit sequence learning. . In M. A. Stadler, & P. A. Frensch (Eds.), The Handbook of implicit learning. (pp. 323).Sage Publications: Thousand Oaks, C A.
- Remmans, A., & McClelland, J.L. (1991). Learning the structure of event sequences. Journal of Experimental Psychology: General, 120, 235-253.
- Wise, L.L. (1993) Melody comprehension and tonal structure. In T. J. Tighe, & W. J. Dowling, (Eds.), Psychology and music: the understanding of melody and rythm. Lawrence Erlbaum Associates: New Jersey.
- Deutsch, J.B. (1978) The Psychology of music. Hutchinson & Co.: London.
- Deutsch, D. (1999) Grouping mechanisms in music. In D. Deutsch, (Ed.), The Psychology of Music (pp.299-300). Academic Press, San Diego.
- Wise, Z. (1993) Compitational models of implicit learning. In D. Berry, & Z. Dienes (Eds.), Implicit learning, theoretical and empirical. (pp. 81). Lawrence Erlbaum Associates, Ltd.

- Dowling, W.J. (1973) The Perception of interleaved melodies. Cognitive Psychology, 5, pp. 322-37.
- Dowling, W.J. (1993) Music cognition and education. In, T.J. Tighe, & W. J. Dowling, (Eds.), Psychology and music the understanding of melody and rhythm (pp.7-9). Lawrence Erlbaum Associates: New Jersey.
- Dowling, W.J. (1999) Development of music perception and cognition. The psychology of music. Academic Press: San Diego
- Fewers, J. H., & Grafel, D. C. (2002) Perception of sonified daily weather records. Preceedings of Human Factors and Ergonomics Society 46th Annual Meeting. 1579-1583.
- Fewers, J. H., & Smith, K. L. (1998). What is learned about nontargets in visual search? Perception and Psychophysics.
- Gallagher, P.A., & Miner, C.S. (1994). Individual differences in short-term memory capacity on an indirect measure of serial learning. Memory and Cognition, 22, 95-110.
- Gelsson, A. (1993) The Complexities of rhythm. In T. J. Tighe, & W. J. Dowling, (Eds.), Psychology and music: the understanding of melody and rhythm. (pp. 93). Lawrence Erlbaum Associates: New Jersey.
- Gorry, A.H., & Varney, N. (1996) Cross-Cultural comparisons in the affective response to music. Psychology of Music, 24, 47-52.
- Herman, J., & Koch, I. (1998) Implicit learning of loosely defined structures. . In M. A. Stadler, & P. A. Frensch (Eds.), The Handbook of implicit learning. (pp. 163).Sage Publications: Thousand Oaks, C A.

- Jackendoff, R., & Lerdahl, F. (1982) A grammatical parallel between music and language. In M. Clynes (Ed.), Music and mind brain, New York: Plenum.
- Mendez, L., & Mendez, C., & Cleeremans, A. (1996). Comparing direct and indirect measures of sequence learning. Journal of Experimental Psychology: Learning, Memory, and Cognition, 22, 948-969.
- Reisberg, M.R. (1993) Dynamics of musical patterns. In T. J. Tighe, & W. J. Dowling, (Eds.), Psychology and music the understanding of melody and rhythm (pp.80-81). Lawrence Erlbaum Associates, New Jersey.
- Lerdahl, F., & Jackendoff, R. (1977) Toward a formal theory of tonal music. Journal of Music Theory, 21, 111-172.
- Wickens, P., & Hill, T., & Bizot, E. (1988). Acquisition of procedural knowledge about a pattern of stimuli that cannot be articulated. Cognitive Psychology, 20, 24-37.
- Miller, R.W. (1967) An objective psychology of music. The Ronald Press Company: New York.
- Wozniak, L., Zizak, D., & Reber, A.S. (1998) Artificial grammar learning and the mere exposure effect. In M. A. Stadler, & P. A. Frensch (Eds.), The Handbook of implicit learning. (pp. 163). Sage Publications: Thousand Oaks, C A.
- Wozniak, 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, 15, 1083-1100.
- Wozniak, L.B. (1994) Emotion and meaning in music. In R. Aiello, & J. A. Sloboda, (Eds), Musical Perceptions (pp31). Oxford University Press: New York.

- e, J.R. (1999) The nature of musical sound. In D. Deutsch, (Ed.), The Psychology of music. (pp.11-12), Academic Press, San Diego.
- r, A. S. (1967). Implicit learning of artificial grammars. Journal of Verbal Learning and Verbal Behavior, 6, 855-863.
- r, A. S. (1969). Transfer of syntactic structure in synthetic languages. Journal of Experimental Psychology, 81, 115-119.
- r, A. S. (1976). Implicit learning of synthetic languages. Journal of Experimental Psychology: Human Learning and Memory, 2, 88-94.
- r, A.S.; Milward, R.B. (1971). Event tracking in probability learning. American Journal of Psychology, 84, 85-99.
- r, A. S.(1993). Implicit Learning and Tacit Knowledge, Oxford University Press: Oxford.
- erson, P. (1989). Verbalizable knowledge and skilled task performance: associations, dissociations and mental models. Journal of Experimental Psychology: Learning, Memory and Cognition, 15, 729-747.
- ks, D.R., & Johnstone, T. (1998) Implicit knowledge in sequential learning tasks. In M. A. Stadler, & P. A. Frensch (Eds.), The Handbook of implicit learning. (pp. 523). Sage Publications: Thousand Oaks, C A.
- h, K.L. (1999) Unintentional Learning and Implicit Performance. Ph.D. Thesis, University of Nebraska - Lincoln
- er, M., & Roediger, H. L. (1996). Question of awareness in research of implicit learning. In M. A. Stadler, & P. A. Frensch (Eds.), The Handbook of Implicit Learning. (pp. 107). Sage Publications: thousand Oaks, C A.

- erwood, G., & Bright, E. H. (1998). Implicit learning. In G. Underwood, (Ed),
Implicit Cognition, (pp. 17) Oxford University Press Inc.: New York
- erwood, G., & Bright, J.E.H. (1998). Cognition with and without awareness. In G.
 Underwood, (Ed.), Implicit Cognition, (pp17). Oxford University Press Inc.: New York,
 1996.
- Noorder, L. (1975) Temporal Coherence in the Perception of Tone Sequences. Ph.D.
 Dissertation. Technishe Hogeschool Eindhoven, The Netherlands.
- at R., Howell, P., & Cross, I. (1985). Musical structure and cognition. Academic Press, Inc.:
 London.

APPENDIX A: ANALYSIS OF THE TASK PERFORMANCE IN

EXPERIMENT 1

AFTER THE 90TH TRIAL1. Score Drops In the Experiment and Nonreversal groups

When analyzing the task performance changes to detect the effects of the reversal, scores for all trials (both same-trials and different-trials) with positive and negative clues were combined. The combined scores were compared for the trial blocks before and after the 90th trial, i.e. for Block 2 (46-90) and Block 3 (91-120). The following table lists the combined mean scores from the two participant groups for those two trial blocks:

Table 26 Means, Standard Deviations and N for Accuracy Scores
of Experiment and Nonreversal groups for Trials before and after 90

	Trial Blocks	
	46-90	91-120
Reversal group	Mean = 0.6284 Std. Dev. = 0.1999 N = 139	Mean = 0.5416 Std. Dev. = 0.2670 N = 139
Nonreversal group	Mean = 0.6018 Std. Dev. = 0.2304 N = 63	Mean = 0.5580 Std. Dev. = 0.2906 N = 63

The change in the mean accuracy score was found to be significant for the reversal group ($t(138) = -3.309$, $p < 0.05$) and but not for the nonreversal group ($t(62) = -1.076$, $p > 0.05$) in which the predictiveness relationships were not changed after the 90th trial.

The change in the mean accuracy score was found to be significant for the reversal group ($t(138) = -3.309, p < 0.05$) and but not for the nonreversal group ($t(62) = -1.076, p > 0.05$) in which the predictiveness relationships were not changed after the 90th trial.

2. Mean Scores of Trials after 90

a) Mean Scores of Same-Trials in the Reversal group

The following table summarizes the summary statistics for same-trials of reversal group:

Table 27 Means, Standard Deviations and N for Accuracy Scores
of Reversal group for Same-Trials after 90

Reversal group: Same-Trials				
Clue	Trial Block	Accuracy Score (Mean)	Accuracy Score (Std. Dev.)	N
Positive	46-90	0.6135	0.1667	35
	91-120	0.5665	0.2168	35
	121-150	0.5582	0.2598	35
Negative	46-90	0.6967	0.2721	35
	91-120	0.5410	0.2004	35
	121-150	0.5432	0.1863	35
Neutral	46-90	0.6912	0.1214	35
	91-120	0.5763	0.3646	35
	121-150	0.5019	0.3398	35

The following figure shows the changes in the mean accuracy scores summarized in the above table:

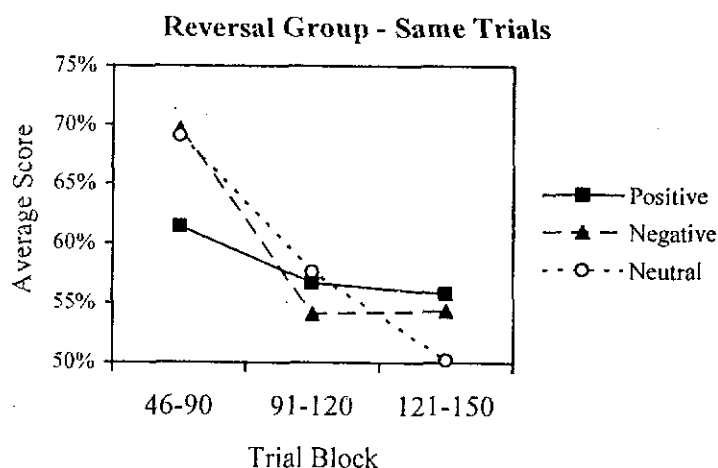


Figure 36 Mean Accuracy Scores of Reversal group for Same-Trials after 90

In the reversal group, for same-trials after 90 (i.e., for blocks 91-120 and 121-150), the repeated-measures ANOVA test did not show a significant main effect of clue type ($F(2,68) = 0.143$, $MSe = 7.843$, $p = 0.867$, $r = 0.21$). Neither was there a significant main effect of trial block ($F(1,34) = 0.599$, $MSe = 6.318$, $p = 0.444$, $r = 0.12$). The interaction between clue type and trial block was not significant ($F(2,68) = 0.406$, $MSe = 7.425$, $p = 0.668$, $r = 0.10$). These results showed that accuracy score of negative and positive clue are not different from each other. In addition, accuracy scores did not change much towards the end neither for negative clue (mean = 0.5410 for trial block 91-120 and mean = 0.5432 for trial block 121-150), nor for the positive clue (mean = 0.5665 for trial block 91-120 and mean = 0.5582 for trial block 121-150). In summary, no significant pattern was found to

indicate that the predictiveness relationships of clues had been re-learned by the reversal-group participants.

b) Mean Scores of Different-Trials in the Reversal group

The following table summarizes the summary statistics for different-trials of the reversal group:

Table 28 Means, Standard Deviations and N for Accuracy Scores
of Reversal group for Different-Trials after 90

Reversal group: Different-Trials				
Clue	Trial Block	Accuracy Score (Mean)	Accuracy Score (Std. Dev.)	N
Positive	46-90	0.5832	0.1960	35
	91-120	0.5546	0.2455	35
	121-150	0.4596	0.2544	35
Negative	46-90	0.6143	0.1399	35
	91-120	0.5068	0.3619	35
	121-150	0.5325	0.3622	35
Neutral	46-90	0.6443	0.2753	35
	91-120	0.4812	0.2654	35
	121-150	0.4973	0.1925	35

The following figure shows the changes in the mean accuracy scores summarized in the above table:

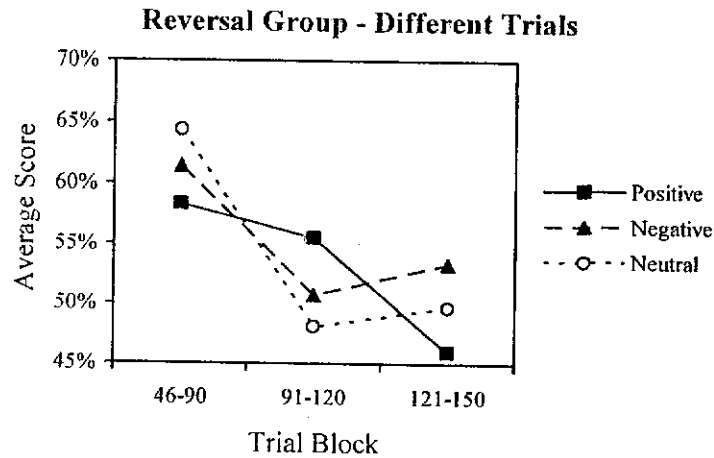


Figure 37 Mean Accuracy Scores of the Reversal group for Different-Trials after 90

For negative clue, mean = 0.5068 for trial block 91-120 and mean = 0.5325 for trial block 121-150. In other words, the mean accuracy score slightly increased towards the end of the session, but since the negative clue was associated with same-trials after the reversal, the mean score for negative-different trials was expected to decrease rather than increase. For the positive clue mean = 0.5546 for trial blocks 91-120 and mean = 0.4596 for trial block 121-150), and that means a decrease, whereas an increase was expected because the reversal of the rules associated the positive clue with different-trials. However, the tests did not mark any of those changes as significant ones. The repeated-measures ANOVA test of different-trials after 90 did not show a significant main effect of clue type ($F(2,68) = 0.227$, $MSe = 7.194$, $p = 0.798$, $r = 0.08$). The main effect of trial block was not significant either

($F(1,34) = 0.338$, $MSe = 4.893$, $p = 0.565$, $r = 0.1$). The interaction between clue type and trial block was not significant ($F(2,68) = 1.095$, $MSe = 7.198$, $p = 0.340$, $r = 0.17$). Thus, no evidence was found for re-learning among the reversal group participants.

c) Mean Scores of Same-Trials in the Nonreversal group

The same analysis was also applied to the nonreversal group scores. The following table presents the summary statistics for same-trials of the nonreversal group of Experiment 1:

Table 29 Means, Standard Deviations and N for Accuracy Scores
of Nonreversal group for Same-Trials after 90

Nonreversal group: Same-Trials				
Clue	Trial Block	Accuracy Score (Mean)	Accuracy Score (Std. Dev.)	N
Positive	46-90	0.7084	0.1834	16
	91-120	0.6394	0.2658	16
	121-150	0.6153	0.3073	16
Negative	46-90	0.5635	0.2555	16
	91-120	0.5417	0.3632	16
	121-150	0.5592	0.4009	16
Neutral	46-90	0.6735	0.1805	16
	91-120	0.6224	0.2850	16
	121-150	0.5494	0.2922	16

The following figure shows the changes in the mean scores listed in the above table:

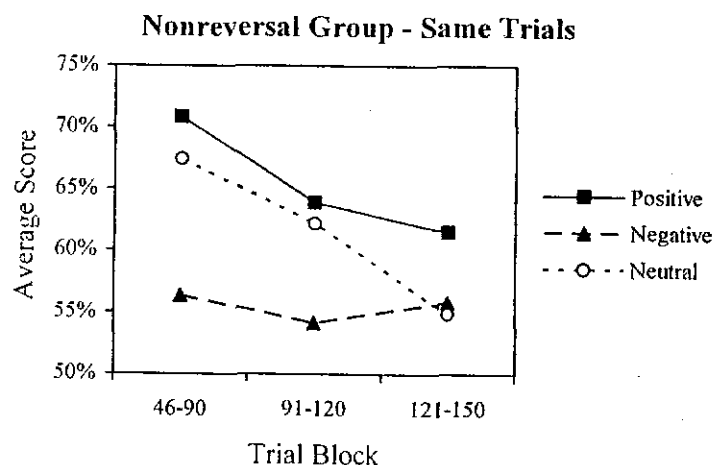


Figure 38 Mean Accuracy Scores of Nonreversal group for Same Trials after 90

The repeated-measures ANOVA test of same-trials of the nonreversal group after 90 did not show a significant main effect of clue type ($F(2,30) = 0.729$, $MSe = 6.515$, $p = 0.491$, $r = 0.21$). The main effect of trial block was not significant ($F(1,15) = 0.223$, $MSe = 7.594$, $p = 0.644$, $r = 0.12$). The interaction between clue type and trial block was not significant ($F(2,30) = 0.171$, $MSe = 9.614$, $p = 0.844$, $r = 0.10$). According to analyses, there is no pattern of any clue learning. These results showed that the mean accuracy score for negative-same trials did not change much towards the end (mean = 0.5417 for trial blocks 91-120 and mean = 0.5592 for trial blocks 121-150). Not much change is observed for the positive-same trials either (mean = 0.6394 for trial block 91-120 and mean = 0.6153 for trial block 121-150). However, the mean score for positive-same trials was higher than

the mean score for negative-same trials, and that was the expected consequence of implicit learning, although that that kind of pattern should have shown itself before the 90th trial.

d) Mean Scores of Different-Trials in the Nonreversal group

The following table presents the summary statistics for different-trials of the nonreversal group:

Table 30 Means, Standard Deviations and N for Accuracy Scores
of Nonreversal group for Different-Trials after 90

Nonreversal group: Different-Trials				
Clue	Trial Block	Accuracy Score (Mean)	Accuracy Score (Std. Dev.)	N
Positive	46-90	0.6105	0.2945	16
	91-120	0.4822	0.2917	16
	121-150	0.5642	0.2244	16
Negative	46-90	0.5853	0.1838	16
	91-120	0.5337	0.2579	16
	121-150	0.5394	0.2559	16
Neutral	46-90	0.5917	0.2884	16
	91-120	0.5417	0.3813	16
	121-150	0.5000	0.3664	16

The following figure shows the changes in the mean score listed in the above table:

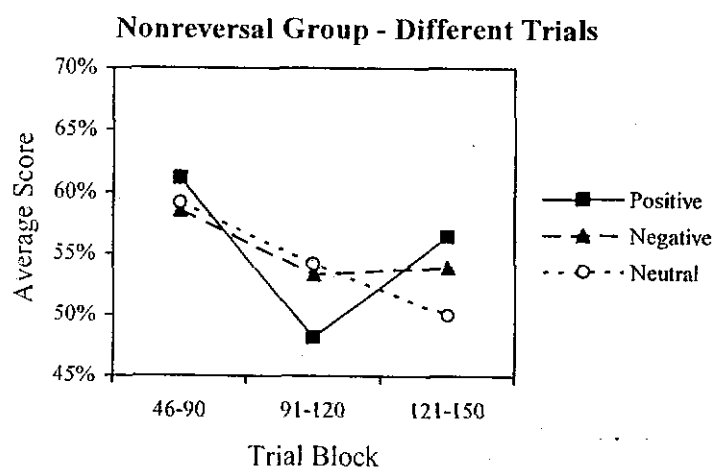


Figure 39 Mean Accuracy Scores of Nonreversal group for Different-Trials after 90

An immediate conclusion is that, the mean score for positive-different trials started to increase towards the end, although it should have remained low. The mean accuracy score for negative clue did not change much towards the end (mean = 0.5337 for trial block 91-120 and mean = 0.5397 for trial block 121-150). The repeated-measures ANOVA test of different-trials of the nonreversal group after 90 did not show a significant main effect of clue type ($F(2,30) = 0.030$, $MSe = 7.635$, $p = 0.971$, $r = 0.04$). The main effect of trial block was not significant either ($F(1,15) = 0.118$, $MSe = 4.773$, $p = 0.736$, $r = 0.08$). The interaction between clue type and trial block was not significant ($F(2,30) = 0.559$, $MSe = 5.570$, $p = 0.578$, $r = 0.18$).

APPENDIX B: MEAN REACTION TIMES OF EXPERIMENT 1

Table 31 Mean Reaction Times of Same-Trials in the Nonreversal group

Clue	Trial Block	Reaction Time (Mean)	Reaction Time (Std. Dev.)
Positive	1-45	6131	1536
	46-90	6323	1638
	91-120	6194	1713
	121-150	6388	1853
Neutral	1-45	6502	1582
	46-90	6341	1916
	91-120	6367	2024
	121-150	5966	1967
Negative	1-45	6173	1720
	46-90	6524	1753
	91-120	6243	1762
	121-150	5978	1774

Table 32 Mean Reaction Times of Same-Trials in the Reversal group

Clue	Trial Block	Reaction Time (Mean)	Reaction Time (Std. Dev.)
Positive	1-45	6189	1483
	46-90	6159	2530
	91-120	5940	1919
	121-150	5560	1980
Neutral	1-45	6144	1557
	46-90	6008	1812
	91-120	5962	2208
	121-150	5349	2194
Negative	1-45	6017	2014
	46-90	6177	1500
	91-120	5897	2340
	121-150	5463	2242

Table 33 Mean Reaction Times of Different-Trials in the Nonreversal group

Clue	Trial Block	Reaction Time (Mean)	Reaction Time (Std. Dev.)
Positive	1-45	6092	1601
	46-90	5917	1758
	91-120	6417	2097
	121-150	6052	1821
Neutral	1-45	6227	1707
	46-90	5993	1725
	91-120	6140	1636
	121-150	5998	2109
Negative	1-45	6279	1761
	46-90	6101	1807
	91-120	6128	1726
	121-150	6095	1808

Table 34 Mean Reaction Times of Different-Trials in the Reversal group

Clue	Trial Block	Reaction Time (Mean)	Reaction Time (Std. Dev.)
Positive	1-45	5879	2281
	46-90	5969	1862
	91-120	5819	1849
	121-150	5601	2244
Neutral	1-45	6058	1995
	46-90	5742	1757
	91-120	5773	1954
	121-150	5419	2157
Negative	1-45	5848	1723
	46-90	5901	1881
	91-120	5509	1767
	121-150	5403	2485

APPENDIX C: MEAN REACTION TIMES OF EXPERIMENT 2

Table 35 Mean Reaction Times of Same-Trials in the Nonreversal group

Clue	Trial Block	Reaction Time (Mean)	Reaction Time (Std. Dev.)
Positive	1-45	5473	972
	46-90	5560	825
	91-120	5458	1026
	121-150	5542	1087
Neutral	1-45	5574	973
	46-90	5453	900
	91-120	5553	1185
	121-150	5355	1277
Negative	1-45	5590	987
	46-90	5525	832
	91-120	5319	1028
	121-150	5411	1164

Table 36 Mean Reaction Times of Same-Trials in the Reversal group

Clue	Trial Block	Reaction Time (Mean)	Reaction Time (Std. Dev.)
Positive	1-45	5811	775
	46-90	5769	915
	91-120	6198	2378
	121-150	5957	1621
Neutral	1-45	5765	1031
	46-90	5891	1335
	91-120	5901	1664
	121-150	5910	1217
Negative	1-45	5781	1007
	46-90	5786	1310
	91-120	5761	1238
	121-150	5823	1683

Table 37 Mean Reaction Times of Different-Trials in the Nonreversal group

Clue	Trial Block	Reaction Time (Mean)	Reaction Time (Std. Dev.)
Positive	1-45	5197	1098
	46-90	4948	1160
	91-120	4725	1195
	121-150	5090	1546
Neutral	1-45	5378	1022
	46-90	4960	1349
	91-120	5140	1401
	121-150	5006	1349
Negative	1-45	4793	1351
	46-90	4900	1458
	91-120	4607	1600
	121-150	4438	1649

Table 38 Mean Reaction Times of Different-Trials in the Reversal group

Clue	Trial Block	Reaction Time (Mean)	Reaction Time (Std. Dev.)
Positive	1-45	5405	1186
	46-90	5724	2638
	91-120	5530	1862
	121-150	5359	2054
Neutral	1-45	5583	1605
	46-90	5415	1991
	91-120	5354	1722
	121-150	5455	1909
Negative	1-45	5328	1120
	46-90	5261	1763
	91-120	5011	1570
	121-150	5380	1481

APPENDIX D: MEAN REACTION TIMES OF EXPERIMENT 3

Table 39 Mean Reaction Times of Same-Trials in the Nonreversal group

Clue	Trial Block	Reaction Time (Mean)	Reaction Time (Std. Dev.)
Positive	1-45	4580	948
	46-90	4506	1255
	91-120	4719	1418
	121-150	4529	1190
Neutral	1-45	4427	730
	46-90	4640	981
	91-120	4598	956
	121-150	4399	1267
Negative	1-45	4677	1291
	46-90	4897	1023
	91-120	4782	1481
	121-150	4509	1394

Table 40 Mean Reaction Times of Same-Trials in the Reversal group

Clue	Trial Block	Reaction Time (Mean)	Reaction Time (Std. Dev.)
Positive	1-45	4542	1437
	46-90	4352	1042
	91-120	4357	1010
	121-150	4305	998
Neutral	1-45	4542	1148
	46-90	4490	1002
	91-120	4397	1071
	121-150	4447	1414
Negative	1-45	4667	934
	46-90	4318	1009
	91-120	4353	978
	121-150	4528	846

Table 41 Mean Reaction Times of Different-Trials in the Nonreversal group

Clue	Trial Block	Reaction Time (Mean)	Reaction Time (Std. Dev.)
Positive	1-45	4446	901
	46-90	4490	1117
	91-120	4359	894
	121-150	4292	1043
Neutral	1-45	4428	919
	46-90	4401	987
	91-120	4327	857
	121-150	4295	1196
Negative	1-45	4573	943
	46-90	4482	1014
	91-120	4526	1113
	121-150	4362	1185

Table 42 Mean Reaction Times of Different-Trials in the Reversal group

Clue	Trial Block	Reaction Time (Mean)	Reaction Time (Std. Dev.)
Positive	1-45	4419	1057
	46-90	4367	932
	91-120	4255	1155
	121-150	4242	1061
Neutral	1-45	4432	1066
	46-90	4393	1086
	91-120	4395	1229
	121-150	4427	1091
Negative	1-45	4339	1188
	46-90	4401	1099
	91-120	4277	1020
	121-150	4378	910

APPENDIX E: MEAN REACTION TIMES OF EXPERIMENT 4

Table 43 Mean Reaction Times of Same-Trials in the Nonreversal group

Clue	Trial Block	Reaction Time (Mean)	Reaction Time (Std. Dev.)
Positive	1-45	4397	720
	46-90	4386	783
	91-120	4293	785
	121-150	4336	702
Neutral	1-45	4480	704
	46-90	4380	731
	91-120	4290	821
	121-150	4242	871
Negative	1-45	4367	881
	46-90	4359	707
	91-120	4343	956
	121-150	4119	895

Table 44 Mean Reaction Times of Same-Trials in the Reversal group

Clue	Trial Block	Reaction Time (Mean)	Reaction Time (Std. Dev.)
Positive	1-45	4350	821
	46-90	4020	775
	91-120	4065	1131
	121-150	3997	985
Neutral	1-45	4268	703
	46-90	4094	882
	91-120	4092	639
	121-150	4158	984
Negative	1-45	4223	679
	46-90	4192	1887
	91-120	3917	876
	121-150	3930	929

Table 45 Mean Reaction Times of Different-Trials in the Nonreversal group

Clue	Trial Block	Reaction Time (Mean)	Reaction Time (Std. Dev.)
Positive	1-45	4094	838
	46-90	4047	979
	91-120	4005	937
	121-150	4003	1128
Neutral	1-45	4368	874
	46-90	3997	829
	91-120	4390	1109
	121-150	3789	1179
Negative	1-45	4063	942
	46-90	3898	973
	91-120	3977	1062
	121-150	3885	991

Table 46 Mean Reaction Times of Different-Trials in the Reversal group

Clue	Trial Block	Reaction Time (Mean)	Reaction Time (Std. Dev.)
Positive	1-45	4043	919
	46-90	3927	984
	91-120	3878	943
	121-150	3660	1011
Neutral	1-45	4122	749
	46-90	4052	1061
	91-120	3882	1161
	121-150	3606	1238
Negative	1-45	3934	877
	46-90	3948	970
	91-120	4008	756
	121-150	3780	1302

APPENDIX F: THE POST-TASK QUESTIONNAIRE

The following is the exact list of questions found in the post-task questionnaire handed to participants of all four experiments.

- 1) As you were performing this task, did you deliberately look for a relationship or a rule which would help you to perform the task faster?
- 2) As you were performing this task, did you notice a relationship or a rule which would help you perform this task of indicating whether two patterns were same or different?
- 3) If you answered "Yes" to the previous question, describe as best as you can the rule or the relationship you made use of.
- 4) Did you find yourself more successful ...
 - a) At the beginning of the experiment
 - b) Towards to middle of the experiment
 - c) At the end of the experiment
- 5) Did you have hard time to stay awake or to keep concentrating?
- 6) If you answered "Yes" to the previous question, describe as best as you can the reasons.
- 7) Did you notice any change the composition or combination of melodies towards at the end of the experiment compared with earlier trial blocks?
- 8) If you answered "Yes" to the previous question, describe as best as you can what kind of a change you noticed.

- 9) Compared to the previous parts, do you think that the very last parts of the experiment was ...
- a) Much easier than the earlier parts of the experiment
 - b) Somewhat easier than earlier parts
 - c) About same level of difficulty as earlier parts
 - d) Somewhat more difficult than earlier parts
 - e) Much more difficult than earlier parts