INFORMATION TO USERS

This material was produced from a microfilm copy of the original document. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the original submitted.

The following explanation of techniques is provided to help you understand markings or patterns which may appear on this reproduction.

1. The sign or “target” for pages apparently lacking from the document photographed is “Missing Page(s)”. If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting thru an image and duplicating adjacent pages to insure you complete continuity.

2. When an image on the film is obliterated with a large round black mark, it is an indication that the photographer suspected that the copy may have moved during exposure and thus cause a blurred image. You will find a good image of the page in the adjacent frame.

3. When a map, drawing or chart, etc., was part of the material being photographed the photographer followed a definite method in “sectioning” the material. It is customary to begin photoing at the upper left hand corner of a large sheet and to continue photoing from left to right in equal sections with a small overlap. If necessary, sectioning is continued again — beginning below the first row and continuing on until complete.

4. The majority of users indicate that the textual content is of greatest value, however, a somewhat higher quality reproduction could be made from “photographs” if essential to the understanding of the dissertation. Silver prints of “photographs” may be ordered at additional charge by writing the Order Department, giving the catalog number, title, author and specific pages you wish reproduced.

5. PLEASE NOTE: Some pages may have indistinct print. Filmed as received.

University Microfilms International
300 North Zeib Road
Ann Arbor, Michigan 48106 USA
St. John’s Road, Tyler’s Green
High Wycombe, Bucks, England HP10 8HR
NISWONGER, Daniel Allen, 1939-
THE DIFFERENTIAL EFFECT OF CONCEPTUAL TEMPO ON THE EFFECTIVENESS OF THE INFORMATION GATHERING STRATEGIES OF SIXTH GRADE BOYS.

The Ohio State University, Ph.D., 1976
Psychology, experimental

© Copyright by
Daniel Allen Niswonger
1976
THE DIFFERENTIAL EFFECT OF CONCEPTUAL TEMPO ON THE EFFECTIVENESS OF THE INFORMATION GATHERING STRATEGIES OF SIXTH GRADE BOYS

DISSERTATION

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

By

Daniel Allen Niswonger, B.S., M.A.

The Ohio State University

1976

Reading Committee:
James K. Duncan
Philip M. Clark
Richard J. Shumway

Approved By

James K. Duncan
Adviser
Department of Curriculum and Foundations
ACKNOWLEDGMENTS

This investigation required the efforts and invaluable assistance of many individuals.

Foremost, my deep appreciation goes to Dr. James K. Duncan for his constant interest, encouragement and guidance through the doctoral program and the present study.

My thanks go to Drs. Philip M. Clark and Richard J. Shumway for their interest and many significant suggestions.

I wish to express my gratitude to Dr. George E. Policello for his assistance with statistical concerns and to Dr. Gerald A. Winer for his suggestions which helped to put the study in perspective.

I also wish to thank Tom and Mary Whitney, Tom for his help with the computer programing and Mary for her assistance in illustrating the 20-questions arrays.

I want to thank my wife, Judy, and my sons, Chip, Steve and Matt, for their help and understanding during the period of my studies. I also want to thank my mother and father, Ruth and Frank Niswonger, for their encouragement.
VITA

July 22, 1939 .................................. Born - Dayton, Ohio
1961 ............................................. B.S., Miami University, Oxford, Ohio
1961-1962 ...................................... Teacher, Jackson Township High School, Preble County, Ohio
1962-1964 ...................................... Teacher, Reading Junior High School, Reading, Ohio
1965 ............................................. M.A. Louisiana State University, Baton Rouge, Louisiana
1965-1966 ...................................... Teacher, Finneytown High School, Cincinnati, Ohio
1966-1969 ...................................... Mathematics Instructor, Miami University, Middletown Branch Campus, Middletown, Ohio
1969-1975 ...................................... Teaching Associate, Mathematics Department, The Ohio State University, Columbus, Ohio

PUBLICATIONS


FIELDS OF STUDY

Studies in Curriculum and Instruction. Professor James K. Duncan
TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>ii</td>
</tr>
<tr>
<td>VITA</td>
<td>iii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>vi</td>
</tr>
<tr>
<td>LIST OF GRAPHS</td>
<td>vii</td>
</tr>
<tr>
<td>LIST OF REPRINTS</td>
<td>vii</td>
</tr>
<tr>
<td>LIST OF SCATTER PLOTS</td>
<td>vii</td>
</tr>
<tr>
<td>Chapter</td>
<td></td>
</tr>
<tr>
<td>I. PROBLEM</td>
<td>1</td>
</tr>
<tr>
<td>Problem and Purpose</td>
<td></td>
</tr>
<tr>
<td>Background of the Problem</td>
<td></td>
</tr>
<tr>
<td>Summary of the Problem</td>
<td></td>
</tr>
<tr>
<td>II. REVIEW OF LITERATURE AND PROBLEM DEVELOPMENT</td>
<td>11</td>
</tr>
<tr>
<td>Conceptual Tempo</td>
<td></td>
</tr>
<tr>
<td>Conceptual Tempo and 20-questions Procedure</td>
<td></td>
</tr>
<tr>
<td>Statement of Hypotheses</td>
<td></td>
</tr>
<tr>
<td>III. PROCEDURE</td>
<td>43</td>
</tr>
<tr>
<td>Description of Pupil Sample</td>
<td></td>
</tr>
<tr>
<td>Data Collection</td>
<td></td>
</tr>
<tr>
<td>Analysis of Data</td>
<td></td>
</tr>
<tr>
<td>IV. RESULTS</td>
<td>61</td>
</tr>
<tr>
<td>E-E'CF-16</td>
<td></td>
</tr>
<tr>
<td>E-NCF-16</td>
<td></td>
</tr>
<tr>
<td>NCF-16 versus E'CF-16</td>
<td></td>
</tr>
<tr>
<td>Conclusions</td>
<td></td>
</tr>
<tr>
<td>Correlational Analysis</td>
<td></td>
</tr>
<tr>
<td>Summary</td>
<td></td>
</tr>
</tbody>
</table>
TABLE OF CONTENTS

ACKNOWLEDGMENTS ............................................... ii
VITA ........................................................................ iii
LIST OF TABLES ...................................................... vi
LIST OF GRAPHS ..................................................... vii
LIST OF REPRINTS .................................................... vii
LIST OF SCATTER PLOTS ........................................ vii

Chapter

I.  PROBLEM ................................................................. 1
Problem and Purpose
Background of the Problem
Summary of the Problem

II.  REVIEW OF LITERATURE AND PROBLEM DEVELOPMENT .......... 11
Conceptual Tempo
Conceptual Tempo and 20-Questions Procedure
Statement of Hypotheses

III. PROCEDURE ............................................................... 43
Description of Pupil Sample
Data Collection
Analysis of Data

IV.  RESULTS ................................................................. 61
E'-E'CF-16
E-NCF-16
NCF-16 versus E'CF-16
Conclusions
Correlational Analysis
Summary
Chapter

V SUMMARY, CONCLUSIONS AND IMPLICATIONS ........................................ 93

Summary
Conclusions
Implications

APPENDIX .................................................. 112

BIBLIOGRAPHY .............................................. 115
# LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>5</td>
</tr>
<tr>
<td>6</td>
</tr>
<tr>
<td>7</td>
</tr>
<tr>
<td>8</td>
</tr>
<tr>
<td>9</td>
</tr>
<tr>
<td>10</td>
</tr>
<tr>
<td>11</td>
</tr>
<tr>
<td>12</td>
</tr>
<tr>
<td>13</td>
</tr>
<tr>
<td>14</td>
</tr>
<tr>
<td>15</td>
</tr>
<tr>
<td>16</td>
</tr>
</tbody>
</table>
### TABLE

<table>
<thead>
<tr>
<th>TABLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>17 COMPARISONS OF HOW EXTREMES (REFLECTIVE-IMPULSIVE)</td>
<td>91</td>
</tr>
<tr>
<td>DISTRIBUTED ABOUT THE MEANS ON TWO MEASURES OF LATENCY</td>
<td></td>
</tr>
</tbody>
</table>

### LIST OF GRAPHS

<table>
<thead>
<tr>
<th>GRAPH</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 E-E'CF-16 EXTREME GROUPS 4 TEMPO GROUPS X</td>
<td>63</td>
</tr>
<tr>
<td>2 ORDERS OF PRESENTATION</td>
<td></td>
</tr>
<tr>
<td>2 E-E'CF-16 ORDINARY GROUPS 4 TEMPO GROUPS X</td>
<td>72</td>
</tr>
<tr>
<td>2 ORDERS OF PRESENTATION</td>
<td></td>
</tr>
<tr>
<td>3 E-NCF-16 EXTREME GROUPS 4 TEMPO GROUPS X</td>
<td>76</td>
</tr>
<tr>
<td>2 ORDERS OF PRESENTATION</td>
<td></td>
</tr>
<tr>
<td>4 E-NCF-16 ORDINARY GROUPS 4 TEMPO GROUPS X</td>
<td>78</td>
</tr>
<tr>
<td>2 ORDERS OF PRESENTATION</td>
<td></td>
</tr>
<tr>
<td>5 E: E'CF-16 vs. NCF-16 (ORDINARY GROUPS)</td>
<td>80</td>
</tr>
<tr>
<td>4 TEMPO GROUPS X 2 GRIDS</td>
<td></td>
</tr>
</tbody>
</table>

### LIST OF SCATTER PLOTS

<table>
<thead>
<tr>
<th>SCATTER PLOTS</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 MFF: LATENCY X ERRORS</td>
<td>49</td>
</tr>
<tr>
<td>2 MFF: LOG (LATENCY) X ERRORS</td>
<td>50</td>
</tr>
<tr>
<td>3 MFF: LATENCY X ERRORS, WITH MEDIAN SPLITS</td>
<td>51</td>
</tr>
<tr>
<td>4 MFF: LATENCY X ERRORS, WITH MEDIAN, 1st AND 3rd QUARTILE SPLITS</td>
<td>54</td>
</tr>
</tbody>
</table>

### REPRINTS OF 20-QUESTIONS ARRAYS

<table>
<thead>
<tr>
<th>ARRAYS</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 E'CF-16</td>
<td>34</td>
</tr>
<tr>
<td>2 NCF-16</td>
<td>39</td>
</tr>
</tbody>
</table>

vii
CHAPTER I

PROBLEM

Introduction

Chapter I provides an overview of the study. It was organized under the following three main sections:

1. Problem and Purpose
2. Background of the Problem
3. Summary of Problem

Problem and Purpose

The purpose of this investigation was to assess the differential effect of conceptual tempo on the effectiveness of the information gathering strategies of sixth grade boys engaged in two variants of the Mosher-Hornsby (1966) 20-questions procedure.

Background of the Problem

Conceptual Tempo

Kagan (Kagan et al., 1964) has empirically demonstrated the pervasiveness of an individual difference variable called conceptual tempo and has identified a response style labeled impulsive-reflective. Kagan's work suggests that the reflective child, when facing a task with high response uncertainty, inhibits his initial response to
consider the differential validity of alternate (potential) solutions, then usually responds correctly. The impulsive child, on the other hand, responds quickly, seemingly reporting the first (potential) solution that occurs to him and is often incorrect.

Siegelman (1969) surmises the distinctions between the cognitively reflective-impulsive child as follows:

The cognitively reflective child is viewed as less likely than the impulsive child to report wrong solutions; more likely to consider alternative possibilities before committing himself; preferring low-risk situations generally, but choosing harder, more solitary intellectual tasks; having a longer attention span; and being less distractable, less motorically active, and more cautious than his impulsive age-mate.

Moreover, Siegelman avowed, and the research of Kagan (Kagan _et al._, 1964) supports, that the characteristics of the reflective dimension appear to be discernible in children (in some form) as early as infancy. Furthermore these differences are pervasive and seem to persist, within the context of an overall trend toward increasing reflectivity with age, and continue throughout adolescence and beyond (Kagan _et al._, 1964).

Conceptual tempo was operationally defined and subjects categorized as slow-accurate (i.e., reflective) versus fast-inaccurate (i.e., impulsive) by applying a split median analysis to the subjects mean latency of initial response and total error score on the Matching Familiar Figures (MFF) test (Kagan _et al._, 1964). In other words, participants scoring above the median on mean latency of initial response and below the median on total errors were classified as having a "reflective" conceptual tempo while respondents scoring below the median on mean latency of initial response and above the median on total errors were labeled "impulsive."
The preceding operational scheme excludes persons scoring at or below the median on both mean latency of initial response and total errors as well as persons scoring at or above the median on these variables. The strong negative correlation between the MFF variables (Messer, 1970) insures that most of the participants will be categorized as fast-inaccurates and slow-accurates with relatively few respondents in the fast-accurate and slow-accurate groups. Nonetheless, the information gathering strategies of the fast-accurates and slow-inaccurates are of interest (Block et al., 1974, Ault et al., 1972, and Eska and Black, 1971) and hence these outlying tempo groups were included in the investigation.

If conceptual tempo has a differential effect on the effectiveness of a respondent's information gathering behavior, perhaps this differential effect is more prevalent at the extremes of the reflective-impulsive dimension. Thus, the investigation involved both "extreme" and "ordinary" tempo groups. The ordinary tempo groups were operationally defined as described on page 2. The extremely reflective and extremely impulsive groups were formed by cutting the MFF errors and mean latency of initial response variables at the first and third quartiles. In short, the extreme reflectives were those respondents who were below the 25th percentile on MFF errors and above the 75th percentile on mean latency of initial response whereas the extreme impulsives were those respondents who were below the 25th percentile on mean latency of initial response and above the 75th percentile on total errors. In line with the strong negative correlation between the
MFF variables, the extreme fast-accurate and extreme slow-inaccurate groups were nonexistent.

20-Questions Task

A number of researchers (Mosher-Hornsby, 1966; Van Horn and Bartz, 1968; Laughlin, Moss and Miller, 1969; Eimas, 1970; Denney, 1973; Ault, 1973; and McKinney, 1973) have used the 20-questions procedure to investigate information gathering-processing strategies.

In the Mosher-Hornsby 20-questions procedure the subjects were shown a 6 x 7 array of familiar objects (See Figure 6, Chapter 3, Studies in Cognitive Growth, Bruner et al., 1966). The participants were then told that the experimenter was thinking of one of the 42 objects and that their task was to guess which object the experimenter had in mind. The respondents were informed that they were to: (1) ask questions which could be answered with a "yes" or "no" response and (2) attain solution in the minimum number of questions.

One approach to the Mosher-Hornsby task was to ask questions of the type: "Is it the airplane?", "Is it the car?" or "Is it the boat?" An alternative line of questioning was: "Is it used in transportation?" or "Is it red?".

Mosher and Hornsby labeled the former strategy specific hypothesis seeking (HS) and called the latter course of action constraint seeking (CS). In short, HS questions encompass but a single object whereas CS questions involve more than one object. A question which has the form of a CS question, but refers to only one object was called a pseudo constraint; e.g., "Does it have a sail?" and the sailboat was the only referent.
Mosher and Hornsby (1966) point out that six year olds use an almost pure hypothesis scanning strategy but by age eleven specific hypothesis questions constitute only a scattering of unresolved temptation. Once persons begin to play the 20-questions game from a constraint seeking approach the issue that arises is: What hierarchical organization will these persons place on their line of questioning?

For instance, an initial constraint seeking question could encompass a few objects (say, 2-4) or it could embody many objects (say 38-40). These situations are complementary. When one categorizes on a few objects a "yes" response effectively reduces the number of possibilities, but the probability of such an occurrence is rather low. On the other hand, when classifying a large number of objects the likelihood of a "yes" response is rather high but, under the circumstances, such a response eliminates few possibilities. Thus, an optimal or safe strategy is one in which the respondent equalizes the usefulness of positive and negative information by asking a sequence of questions that repeatedly encompass half (or as near half as possible) of the remaining candidates.

Eimas (1970) and McKinney (1973) employed a formula from Shannon's (1948) mathematical theory of communication to assess how closely the hierarchical organization of a respondent's line of questioning corresponds to the optimal "half by half ..." principle. The formula is:

$$E_i = -[p_i \log_2 p_i + (1 - p_i) \log_2 (1 - p_i)]$$

where $E_i$ is the expected informational value (i.e., the expected reduction in uncertainty--measured in bits) of the $i^{th}$ question of a given sequence of questions and $p_i$ is the probability of a "yes" response to the $i^{th}$ question.  

1A development and discussion of $E_i$ scores can be found in Lee's (1971, p. 264) Decision Theory and Human Behavior or in Berlyne's (1960, p. 25) Conflict, Arousal and Curiosity.
Perhaps the computation of some selected $E_i$ scores would be instructive. First, suppose a respondent asks a question that classifies half of the candidates. Thus,

$$E_1 = -\left[0.5 \log_2 0.5 + 0.5 \log_2 0.5\right]$$

$$= -\left[0.5(-1.0) + 0.5(-1.0)\right]$$

$$= -[-1.0] = 1.0$$

Second, suppose a participant asks a question that categorizes one fourth of the candidates. Hence,

$$E_2 = -\left[0.25 \log_2 0.25 + 0.75 \log_2 0.75\right]$$

$$= -\left[0.25(-2.000) + 0.75(-0.415)\right]$$

$$= -[-0.811] = 0.811.$$  

Third, suppose a subject classifies one eighth of the remaining candidates. Whence,

$$E_3 = -\left[0.125 \log_2 0.125 + 0.875 \log_2 0.875\right]$$

$$= -\left[0.125(-3.000) + 0.875(-0.193)\right]$$

$$= -[-0.544] = 0.544.$$  

Thus, if a respondent eliminates half of the alternatives they obtain an $E_i$ score of 1.0 and the more a participant's questions deviate from the "half by half..." principle the lower their $E_i$ score. Furthermore, it is obvious that the direction of the deviation will not alter the $E_i$ score; e.g., if questions two and three had classified three fourths and seven eighths of the candidates, rather than one fourth and one eighth, the respective $E_i$ scores would have
remained the same. Neimark and Lewis (1967), Eimas (1970) and McKinney (1973) obtained a mean expected information score for each subject by determining the $E_i$ score for each question of a sequence of questions and then summed and divided by the number of questions in the sequence.

The decision to use $E_i$ scores as a means of assessing the effectiveness (i.e., hierarchical arrangement of questions) of a subject's information gathering strategy requires that the questions unambiguously partition the array and thereby provide for the objective assignment of probabilities to each partition class. The Mosher-Hornsby array does not satisfy this proviso. For example, the query, "Is it a toy?" leaves one undecided as to whether the sailboat is a toy or a real sloop, and hence, one cannot objectively assign probabilities to the partition classes induced by this question.

The arrays of Eimas (1970) and McKinney (1973) contained rather simple color-form attributes and therefore, on these arrays, there is (virtually) no difficulty with ambiguous classification. The Eimas and McKinney arrays are 16-item 4 by 4 grids with four equally distributed bivariate dimensions. In other words, each of the sixteen distinct objects consists of four attributes with each attribute fixed on exactly one of two levels. For example, the Eimas color-form array consisted of: shapes, diamond vs. square; color, red vs. blue; letter, S vs. 0 and size of letter, big vs. small—in all there were eight diamonds and eight squares, eight reds and eight blues and so on.

Since the Eimas and McKinney arrays were constructed on this convenient $2^4$ basis, any question which categorized a single attribute at a given level automatically partitioned the remaining candidates
into numerically equal subsets; hence, one can acquire the optimal strategy quite incidentally. At any rate, it seems that the $2^4$ proportioning would facilitate the use of the optimal "half by half. . ." strategy. Thus, the proportioning of array attributes may have a confounding effect on the hierarchical organization of a respondent's line of questioning. In order to test the influence that proportionality of attributes exerts on the effectiveness of a respondent's information gathering strategy (i.e., the hierarchical organization of their line of questioning) the investigator included two rather similar color-form arrays; one array proportioned on the $2^4$ basis and the other array lacking the $2^4$ proportioning.

Since the array with the $2^4$ proportioning has been alleged the "easier" of the two arrays, it was of interest to investigate the possibility of an order of presentation effect. One might predict that presenting the "easier" array first would facilitate optimal solution on the more "difficult" array. However, the different proportioning of the two arrays alters the nature of the problem and hence, one might predict that, if anything, presentation of either array first would have a negative influence on the second presentation. Although Eimas (1970) reports no "order of presentation" effect, there is the possibility of an order of presentation-conceptual tempo interaction.

**Summary of Problem**

Reflectives were characterized as less risk taking, more cautious and more apt to consider the validity of potential solutions than their impulsive age-mates. Thus, reflective respondents would be
more likely than their conceptual tempo counterparts, to consider the consequences of their 20-questions queries and thereby employ a line of questioning in which the hierarchical organization more closely approximates the optimal "half by half. . ." principle. In short, one would predict that the reflective subjects $E$ scores would be significantly higher than those of the other tempo groups.

The empirical testing of the aforementioned prediction may be confounded by the proportioning of the dimensions of particular 20-questions arrays. This confounding was controlled by employing two 20-questions arrays. The one 16-item color-form array had the $2^4$ proportioning which (it was predicted) would facilitate the use of the "half by half. . ." principle while the other 16-item color-form array lacked the convenient $2^4$ proportioning. Thus, besides testing the conceptual tempo prediction for both arrays, the prediction that the $2^4$ proportioned array will be easier (i.e., the $2^4$ proportioned array will post significantly higher $E$ scores) will also be tested.

Furthermore, the conceptual tempo predictions (for both arrays) were tested for both "ordinary" and "extreme" tempo groups. Although an extreme groups design has seen little or no use in conceptual tempo research, one would expect any differential effect due to conceptual tempo to be more prevalent at the extremes, and hence, the nonuse of an extreme groups design increases the likelihood of overlooking significant differences. Since the ordinary tempo groups were the most representative of the population at large, the array difficulty prediction was tested only for ordinary tempo groups.
Lastly, the possibility of an order of presentation effect as well as the possibility of an order of presentation-conceptual tempo interaction were tested, separately, for each of the two arrays. These tests were run for the purpose of detecting an order of presentation or interaction effect rather than for the purpose of testing a particular directed hypothesis.
CHAPTER II
REVIEW OF LITERATURE AND PROBLEM DEVELOPMENT

Introduction

Chapter II contains a review of the conceptual tempo and 20-questions game literature with problem development and a statement of hypotheses as focal points. Chapter II was organized under the following main sections:

(1) Conceptual Tempo
(2) Conceptual Tempo and 20-Questions Procedure
(3) Statement of Hypotheses.

Conceptual Tempo

Conceptual Tempo and Problem Solving

Kagan et al. (1964) offers the following view of the relationship between problem solving and conceptual tempo.

In complex problems with alternate routes to solution, reflection upon the probable validity of a varied solution sequence is critical for the ease with which success is achieved. The child who does not reflect upon the differential validity of several solution possibilities is apt to implement mentally the first idea that occurs to him. This strategy is more likely to end up in failure than one that is characterized by reflection. For the impulsive child who reaches a cul-de-sac in a problem solving sequence and recognizes he has not solved the problem is likely to become more anxious than he was initially. As a result of the increased anxiety, his selection and evaluation of a second solution path is apt to be impaired and the probability of success attenuated.
Milgram (1969) found the correlation between the combined scores of all subtests on the Forstag Visual Perception Test and the classifications (i.e., impulsive-reflective) resulting from the Matching Familiar Figures Test (MFF) to be .77 (p < .05 two tailed). This rather high correlation between the reflective-impulsivity dimension of cognitive style and visual perception (for 5 year olds) has led Milgram to assert that, "Further evidence is needed to assume that the dimension of cognitive style which the MFF purports to measure is actually related to reflective and impulsive thought processes."

Klausmeir and Laughlin (1961) stratified a population of 5th graders into three groups (high, medium and low) on the basis of (WISC) IQ scores and found that, when facing problems\(^1\) graded in difficulty to match their present achievement levels, subjects of the highest IQ strata showed a greater incidence of persistence, correcting mistakes, independently verifying solutions, and utilization of a logical approach than subjects of the middle IQ groups. Moreover, the above trend was exhibited across the three strata with the lowest IQ group showing the greatest incidence of non-persistence, offering the most incorrect (purported) solutions without attempted verification, and utilizing a random rather than logical approach. However, Klausmeir and Laughlin also reported that differences in approaches and performances among individuals within the various IQ strata were large.

\(^1\)Klausmeir and Laughlin posed problems in which subjects were given fixed dollar and cents amounts as well as a prescribed number of monetary pieces; subjects were then required to choose monetary pieces of various values in an attempt to attain the fixed amount by using (exactly) the allotted number of monies.
The issue raised by Milgram—perhaps (cognitive) style of functioning on the MFF may be specific to some attribute of the task at hand (say, visual feature analysis) rather than generalizable to one's mode of functioning and "thinking" on a wide range of problem solving tasks—is a question that cannot be ignored; furthermore, the rather strong relationship between conceptual tempo and the Forstag Test of Visual Perception supports Milgram's view. However, the findings by Klausmeir and Laughlin (on the coin problem) illustrate the existence of differential cognitive styles on a presumably nonvisual task; it further appears that these differences are not strictly related to one's score on the WISC Intelligence Test.

Kagan et al. (1966) demonstrated that impulsivity adversely affects performance on inductive reasoning tasks. An examination of the details and results of this inductive reasoning study will be undertaken to (1) show that tasks of a highly visual nature can involve a substantial verbal component and (2) to illustrate that the benefits of a preference for reflective delay pervades across nonvisual tasks.

The inductive reasoning tasks (Kagan et al., 1966) included a Picture Completion Reasoning test and a Guessing Objects test. The picture-completion items contained an ordered sequence of three pictures that told the beginning of a story. Four more pictures were then presented and the respondent was asked to select the one picture from the four that was the next thing that happened in the story.

The extrapolation reasoning items consisted of a series of visual stimuli that were linked by a principle; e.g., big rectangle, small rectangle, big triangle, small triangle, big square, small square and
big circle. The participant was also shown a second set of visual stimuli and asked to select the one that continued the preceding principle; i.e., small circle.

In the guessing task the respondent was verbally given three characteristics of an object and then asked to guess the object; e.g., "what is round, small, and found on a shirt?"

The respective correlations between MFF response time and errors on picture-completion, extrapolation and guessing tests were - .25, - .40 and - .35 for boys and - .45, - .29 and - .29 for girls. Likewise, the respective correlations between MFF errors and errors on the above inductive reasoning tasks were .25, .46 and .27 for boys and .37, .38 and .15 for girls. (All coefficients larger than .22 were significant at p < .05 for two tails).

The aforementioned findings indicate that a reflective disposition was associated with accurate performance on the inductive reasoning tasks. For boys, the relationship between a reflective disposition and accurate performance on the inductive reasoning tasks, was strongest for the extrapolation task whereas the strongest relationship for females was between a preference for reflective delay and accurate performance on the picture-completion task. Observe, however, that both the extrapolation and the picture-completion tasks entail visual perception. On the other hand, it appears that (at least for boys) a reflective disposition was associated with accurate performance on the (nonvisual perception) guessing task.
In the same study the respective correlations (Kagan et al., 1966) between WISC-verbal-IQ and errors on the picture-completion, extrapolation and guessing tasks were -.56, -.44 and -.27 for boys and -.27, -.43 and -.17 for girls. Moreover, neither of the MFF variables (response time or errors) were significantly associated with WISC-verbal-IQ; however, the correlations (both boys and girls) between MFF errors and WISC-verbal-IQ just missed significance at p < .05 for two tails.

Although the extrapolation task requires recognizing and continuing a principle that is embedded in a sequence of visual stimuli, it appears that for boys and girls alike, being highly verbal is associated with accurate performance on this reasoning task. Maybe verbally labeling the components of the visual sequence (e.g., big rectangle, small rectangle and so on . . .) facilitates the quality of one's performance. In short, this verbal labeling may act as a mediating device and, perhaps, the highly verbal children are more apt to employ such a procedure.

More importantly, there could be an interaction between visually analyzing the sequence into components and the attaching of verbal descriptors to these components. In other words, maximal proficiency may be the result of a combined visual-verbal process—even on tasks that appear primarily visual or primarily verbal. Perhaps, an interactive process of visual analysis-verbal descriptors may require a period of reflective delay.

Moreover, of the three inductive reasoning tests, it is surprising that success on the guessing task—a task in which the necessary
information was transmitted by strictly verbal means—bares the least relationship with (WISC) verbal ability. In fact, accurate performance on the guessing task bares at least as much or more association with lengthy MFF response times than with WISC-verbal-IQ. At any rate it appears that hasty dichotomization of tasks as strictly visual-nonvisual or strictly verbal-nonverbal may prove wrongheaded.

A pertinent issue has been submerged beneath the surface of the foregoing discussion. The questions that arise from this issue are: Is it the case that impulsives are "faster thinkers" than reflectives but otherwise they cognitively function in much the same fashion? Or, is it the case that the reflective's delay is primarily due to differential cognitive activity? Perhaps the next two sections will shed some light on the above issue.

**MFF Strategies**

In several investigations (Siegelman, 1969; Nelson, 1969; Drake, 1970; Zelniker et al., 1972 and Ault et al., 1972) mechanical devices were employed to study the looking behavior and eye fixations of participants engaged in MFF type tasks. In these studies the investigators inferred MFF strategies from the recorded visual fixations. These studies, however, have produced some conflicting results. The conflicts arose when relative observing responses were considered in lieu of absolute measures. The controversy centers around the use of a standard to variant paired comparison MFF scanning strategy versus a strategy of comparing across variants (presumably to locate differences) and then checking the standard to eliminate the deviants.
Siegelman (1969) and Nelson (1969) report that reflective fourth-grade boys devote proportionally less time to the standard and correspondingly more time to the variants, than do their impulsive peers. Drake (1970) conducted a bisexual study and obtained results similar to the above for college age respondents but obtained opposite results for third grade children. Zelniker et al., (1972) and Ault et al., (1972) failed to find any such differences for third grade children. Ault et al. (1972) included fast-accurate and slow-inaccurate children in their study. Moreover, these researchers report that although the children relied heavily on a basic standard to variant paired-comparison MFF scanning strategy, the accurate children, whether slow or fast, tended to be more systematic in the use of that strategy than their inaccurate (slow or fast) counterparts.

On the grounds of the above, Block et al. (1974) remarks, "There is nothing that can be adduced from these several studies on scanning strategies"; one should, however, differentiate between the specific executorial aspects of a participant's MFF scanning strategy and their overall (i.e., general game plan) MFF strategy. Two diametrically opposite "overall" MFF strategies can be characterized by the way a respondent defines a "match" between a variant and the standard. First, a participant may declare a "match" between a variant and the standard if they look like they match. Secondly, a subject may declare a variant and the standard "matched" if and only if all other variants have been eliminated. In the first case the participant places a heavy reliance on their immediate perceptions while the second case indicates a mistrust of one's immediate perceptions and a reliance on a logical process or
algorithm for matching. Furthermore, the two positions are indicative of the stock one places in the "right to be sure" (Ayer, 1956).

It appears that behaving in adherence to the "match on a look-a-like basis" as opposed to a thorough and systematic "eliminative procedure" would be crucial to one's MFF outcomes. In short, it is suspected that the more impulsive subjects are matching (primarily) on a look-a-like basis whereas the more reflective subjects are matching by a more thorough and systematic process of elimination.

Moreover, there is no strict one-to-one correspondence between a participant's overall matching strategy and the aforementioned controversial scanning strategies. For instance, a respondent who attempts to find the MFF match by carrying out a thorough and systematic search and elimination procedure can conduct the search by making repeated pairwise comparisons between standard and variants or by comparing across variants to locate differences and then checking the standard to eliminate deviants. Perhaps the use of a scanning across variants and then checking the standard approach instead of a variant to standard paired comparison strategy is more indicative of a high level of executional sophistication rather than being indicative of a respondent's overall matching strategy.

If one abandons the more executional aspects of MFF strategies, is there any evidence indicating that reflective subjects display a preference for a thorough, systematic process of elimination while their impulsive age-mates show a preference for matching on a look-a-like basis? Siegelman (1969) and Drake (1970) report that reflective subjects evidenced a less peaked and more even distribution
of attention across the MFF variants and more frequent comparisons among homologous parts of the MFF stimuli than their impulsive peers. Zelniker et al. (1972) reports that prior to initial response, reflective subjects fixated on significantly more of the MFF variants than their impulsive counterparts. Ault (1972) remarks, "The Fast/Accurate Ss are not eliminating all of the incorrect variants prior to responding. Rather they are seeking to match the standard by making comparisons between the standard and some of the variants."

In short, when one considers the totality of the above studies, they suggest that reflective subjects employ a rather thorough and systematic search and elimination procedure whereas the other tempo groups appear to falter (in varying degrees, of course) in the use of such a process. This view concurs with the findings of Usui (1975).

In all, it appears that the different conceptual tempo groups are marked by differential cognitive activity as well as by different temporal expenditures. Perhaps the participants who opt for a rather thorough and systematic search and elimination procedure place less faith in their immediate perceptions. At any rate, the relationship between the differential cognitive processes of the various conceptual tempo groups (namely, reflective-impulsive) and their subsequent time expenditures will be further explored in the next section.

**Attempts to Modify MFF Performance**

A number of researchers have investigated training procedures aimed at lengthening latency of response on MFF items. Some of these investigators (Briggs, 1966; Kagan et al., 1964 and Messer, 1968) report
that improvements in the accuracy of MFF performance accompanied the extended response times while others (Debus, 1970; Denney, 1972 and Kagan, et al., 1966) report increased latency with no accompanying effects on accuracy. Moreover, in order to improve MFF accuracy consistently, it has been necessary (Meichenbaum and Goodman, 1971; Nelson, 1968; Hedier, 1971 and Zelniker, et al., 1972) to use training procedures directed toward the implementation of more effective search and elimination strategies; i.e., "process" training rather than latency training per se.

These studies will be examined in greater detail. Both Messer (1968) and Kagan, et al. (1966) manipulated "anxiety" in an attempt to increase response latency and improve MFF performance. Kagan et al. (1966) modeled adult anxiety over errors while Messer (1968) aroused general anxiety over intellectual performance, and, although both investigators were successful in their attempts to lengthen response latency, only Messer recorded a corresponding improvement in MFF accuracy.

Schwebel (1966) enforced a delayed response on a sentence construction task and found that the enforced "time lag" significantly improved performance. Heider (1971) replicated Schwebel's study and reported similar findings; in addition, she found that a direct "process" instruction condition yielded a more pronounced improvement than the "time lag" condition.²

²The Schwebel and Heider samples were restricted to caucasian males from lower and middle socio-economic classes. The participants from the lower socio-economic class were significantly more impulsive than their middle class counterparts. Furthermore, the improved performance findings pertain only to the lower class groups. The lack of improvement for the middle class participants appears to be due to a "ceiling effect."
Heider's direct "process" instructions were, "Say your sentence to yourself before responding." It may be that both the delayed response and the direct "process" instruction conditions evoke (the delayed response condition to a lesser degree, of course) the "rehearsal process." Perhaps, on tasks in which accuracy of performance is associated with long response latencies, a forced delay is a sufficient condition for increased accuracy if and only if the forced delay is a sufficient condition for envoxing a "proper" strategy.

Zelniker et al., (1972) also employed an intervening "process" treatment and reports that the modified scanning strategy found in the treatment task transferred to post MFF problems for impulsives only. Moreover, these impulsives evidenced a substantial decline in MFF errors with no accompanying change in MFF latency. Even though the experimental impulsives were the only group to significantly improve their MFF accuracy on task III, their task III errors still remained substantially greater than that of the experimental reflectives.

Bush and Dweck (1975) presented one hundred fourth graders (reflective-impulsive by high-low anxious) with speeded tasks of increasing difficulty. Contrary to prediction, the results revealed that high-anxious reflectives performed as well as low-anxious reflectives and both were generally faster and more accurate than impulsives.

Denney (1973) found that on the Mosher-Hornsby 20-questions array 7-8 year old reflective subjects ask significantly more constraint seeking (CS) questions than their impulsive age-mates. The participants

3 Task II, the treatment task, consisted of having subjects find the familiar figure that was different rather than matching familiar figures.
of Denney's (1973) study were also instructed to hasten or delay their responses on a test of specific hypothesis seeking (HS) versus constraint seeking (CS) conceptual strategies. Data on pre-training and immediate post tests revealed that the desired alterations of response latency were successful. Furthermore, subjects who were instructed to hasten their responses were found to significantly increase their use of HS strategies while subjects who were asked to delay their responses showed no significant change in strategy.

White's (1965) temporal stacking model suggests that following the presentation of a stimulus, different responses (qualitatively speaking) are maximally available during different temporal zones—with the more sophisticated response lying in a latter zone. The results of Denney's (1973) investigation support the contention that "time lag" is a necessary but not sufficient condition for the emergence of the more sophisticated CS response. This concurs with the general findings on "time lag" and points to the position that, perhaps, reflectives are not simply slower and more meticulous at doing what impulsives do rapidly, but are indeed going about the task by a more elaborate and more sophisticated "process."

Overall, the studies on the modifiability of conceptual tempo—particularly the investigation by Zelniker et al. (1972), Denney (1973) and the findings of Bush and Dweck (1975)—are suggestive and supportive of the position that task performance by reflective-impulsive subjects may be more a function of the strategy used rather than long or short response times per se. Since the reflective delay appears to be due to differences in cognitive processes rather than slowness per se, it is
plausible, to conjecture that the disposition for this differential cognitive activity is generic in nature and thereby pervades across a broad range of tasks which entail alternative courses of action. Thus, a survey of the research which attempts to relate MFF performance to other tasks will now be undertaken.

MFF and Other Tasks

In this section an examination of the task conditions which are essential to achieve maximal benefit from a preference for reflective delay \(^4\) will be undertaken. This examination will focus on two main issues:

1. the effect of task difficulty \(^5\) on the benefits of a preference for reflective delay;

2. the effect that the task potential for \(a\ priori\) "correct" responses can have on the benefits of a preference for reflective delay.

Impulsivity has been shown to adversely affect reading performance (Kagan, 1965b and 1965c). In Kagan's study subjects were given

\(^4\) Reflective delay does not refer to delay per se, but indicates time lapse due to the employment of a more systematic, thorough and thereby more time consuming, search and elimination procedure.

\(^5\) Difficulty as used herein refers to the degree of fluency to which the participants have acquired and accomplished the rudiments of the task.
a word recognition task that involved items in which a word such as "moon" was read by E. Then S was shown a card containing words such as noon, moon, boom, soon and mean and was asked to point out the correct word. Since the word recognition task requires visual discrimination, as does reading in general, the relationship between MFF performance and reading performance comes as no surprise. On the other hand, Kagan's remark (1965b and 1965c) that "the influence of reflective delay is maximal when the subject has already learned the rudiments of the skill necessary to perform a task but has not over learned the skill" may prove rather illuminating.

For example, Kagan reports (1965b and 1965c) that among high-verbal children (first graders assessed by the WISC) the correlations between response time on MFF and word recognition errors were -.40 for boys and -.44 for girls while corresponding correlations for low-verbal children were -.14 for boys and -.21 for girls. It appears that the high-verbal children have acquired the rudiments of word recognition to the extent that reflection over the validity of alternative choices (perhaps an elimination process) benefits the quality of their performance. For low-verbal children, children who perhaps have difficulty recognizing letters or children who perhaps have not yet formed adequate associations between letters and their respective sounds, it appears that reflective delay is not substantially related to accurate word recognition.

Thus, in studies on the effect of the reflective-impulsive dimension, task selection is of prime concern. In short, the task can not be overwhelmingly difficult nor can it be so easy that uncertainty is severely attenuated.
In another study Adams (1972) used the Weir Marble Game (1964) apparatus to investigate the problem solving strategies of reflective-impulsive subjects. The Weir Marble Game apparatus consisted of a panel containing a horizontal row of three knobs. Above the knobs was a signal light and below the knobs was a delivery hole for the marbles. Each subject was, individually, seated in front of the apparatus and was told that when the signal light went on they were to press one of the three knobs. They were also told that if they were correct, a marble would fall into the enclosed container. All subjects were informed that the object of the game was for them to win as many marbles as possible. For each subject, a knob was selected as their "correct" knob and this "correct" knob was set on a 33 per cent random reward schedule; i.e., S is randomly—rather than every third time—rewarded 33 per cent of the time when they choose the knob selected as correct for them. Also, each subject was given 80 trials—eight blocks of ten trials.

Adams (1972) reports that impulsive subjects made significantly more "correct" knob responses than the reflective subjects. Further analysis of the interaction components revealed that young impulsives (six year olds) made significantly more "correct" responses than six year old reflectives, eight year old impulsives and eight year old reflectives. These latter groups did not significantly differ.

The direction of these findings may seem somewhat surprising; however, these results are basically in agreement with the findings of Weir (1964). Weir avows, "the belief by older subjects that there's
a complex solution actually results in fewer choices of the most frequently reinforced alternative." The point that subjects who tend to expect complex solutions are hampered in some problem solving tasks has been noted by Kendler (1963); Stevenson, Iscoe and McConnell (1955); Weir and Stevenson (1959). Also, in line with this view is Hyman and Jenkins (1956) finding that it is much more difficult to convince adult subjects that a sequence is random than it is to convince these subjects that it is structured.

At this point a look at the nature of the marble game is in order. First, the marble game is such that one can not arrive \textit{a priori} at a well-founded "correct" rule or strategy. In other words, a marble game player must make some initial pulls on the various knobs, generate some data, and then, on the basis of these data, attempt to attain a valid rule or strategy for maximizing the collection of rewards. Or, if worse comes to worst, one may conclude that the data are insufficient and therefore a repetition of the above process is in order. Further, there is no valid reason why a rule that appears to satisfy the data of the first fifteen trials must also apply to the next fifteen trials. Hence, if on the basis of past evidence, one pieces together a rule or strategy for maximizing rewards, then one has no "right to be sure" whether the said rule is for the full range of the marble game or not.

Secondly, the notion of an "optimal" or "correct" strategy for the marble game is a bit of a misnomer. E knows that there is a "correct" strategy and furthermore E knows what it is. However, S can only take two or more strategies, and via the generated data, decide
which of these strategies appears to be the better or best; but S can never be sure that this "better" or "best" strategy is "the optimal" strategy.

Now, as the marble game is currently played, if one is to be maximally successful in the collection of rewards, it is necessary to choose the "correct" knob on the initial move and then persist in this choice throughout the game. Moreover, it becomes clear that to conduct an extensive, thorough, and on-going research project in an attempt to discover that such and such is the "good" knob, or that such and such is the (or a) "proper" rule or strategy for playing the game is, by and large, not the way to acquire the most rewards--especially in a "short run" situation. Hence, "success" in the marble game depends on one's being able to arrive at a "correct" conclusion or choice of knobs rather quickly; possibly at the expense that the evidence on which the decision is based may be skimpy and, indeed, leave one open to the accusation of "jumping to conclusions." Thus, it seems, that if a task contains the room or the potential for participants to distinguish good moves from poor moves prior to their initial response, then the benefits from a preference for reflective delay would be enhanced.

Kochen and Galanter (1958) on a data collection task which was not unlike the marble game, handled the above problem by classifying a participant's responses as "investigative or information-gathering" versus "game playing or playing to win." This type of classification indicates an awareness of the fact that one can approach the marble game from the position "experiment-attempt to get the rule-then play
the game" or, alternatively, one can approach the game from the view
"play to win--from the start."

Because of the nature of the marble game, Weir (1964) and Adams
(1972) use this task (primarily) to investigate problem solving stra-
tegies. Weir's (1964) norms as well as the previously cited remarks
on page 26 suggest that older subjects are more apt to employ the
"experimental" approach to the marble game while younger subjects tend
to play to win from the start. Moreover, Adams (1972) reports that
if one collapses the tempo groups and simply considers six year olds
versus eight year olds then his younger subjects performed like Weir's
5-6 year old group while his older respondents performed like Weir's
7-9 year olds. However, when one groups on both tempo and age, Adams
reports that the younger impulsives functioned like Weir's 5-6 year olds,
both of the older tempo groups performed like Weir's 7-9 year olds while
the younger reflectives evidenced behavior which was similar to that
of the 7-9 year olds.

Adams further points out that, although the actions of the young
reflectives tend to mock the behavior of the older participants, the
older subjects appear to give up the right-middle-left or left-middle-
right patterning as nonproductive whereas the younger reflectives seem
to persist with this nonproductive strategy. These findings concur
with Elkind's (1958) characterization of concrete operational
thought as the ability to generate hypotheses but not the ability to
evaluate effectiveness in light of the evidence. In short, the effec-

---

6 Adams (1972) reports that young impulsives never used the
patterned responding until the very end of the task.
tiveness and results that younger reflectives obtain from their more mature behavior seem to be confined and restricted by the developmental and intellectual constraints of a younger age group.

Although Adams older reflectives evidenced a marked "correct" choice improvement on the eighth block, it appears that at the older age level conceptual tempo has little or no differential effect on the Adams-Weir version of the marble game. Perhaps the findings with the older group are due to the fact (Weir's norms for 80 trials-1964) that between 7-9 years of age variability is greatly constructed; it may be that more than 80 trials are necessary to produce substantial variability at the 7-9 year age level.

The marble game findings acquire an added note of importance when viewed in light of the fact that this task requires an analysis of action (i.e., the pulling of knobs and then the occurrence of a reward or nonreward condition) whereas the MFF task entails, among other things, visual feature analysis.

Conceptual Tempo and 20-Questions Task

Review of Tempo and 20-Questions Studies

A number of investigators (Denney, 1973; Ault, 1973 and McKinney, 1973) have studied the relationship between a respondent's conceptual tempo and their 20-questions performance. The studies of Denney (1973) and Ault (1973) were primarily concerned with the sophistication of a subject's questions; e.g., the use of specific hypothesis seeking (HS) questions versus the use of constraint seeking (CS) questions. Moreover, Denney (1973) and Ault (1973) report that although reflective
subjects asked significantly more sophisticated questions than their impulsive peers, they were no more effective in attaining solution.

The finding that the use of a more sophisticated strategy did not lead to a more efficient solution is not as puzzling as it might initially appear. Quite simply, in the 20-question game, question sophistication and question effectiveness are not one and the same thing. For instance, when 20-question arrays contain more than three candidates, any optimally effective initial question must be a constraint seeking question, and hence, the question contains some degree of sophistication. However, as seen in Chapter I (p. 5), not all constraint seeking questions have the same degree of effectiveness; e.g., constraint seeking questions which classify all but one grid object are no more effective than their complementary specific hypothesis seeking counterparts. In short the use of sophisticated questions does not insure that the hierarchical organization of the subject's line of questioning conforms to the optimally effective "half by half..." principle.

Neither Denney (1973) nor Ault (1973) assessed 20-questions effectiveness by use of E scores. (See page 5). Denney assessed 20-questions effectiveness by the number of questions required for solution—with the less effective subjects being those who required more questions. Ault assessed 20-questions effectiveness by counting the number of objects the subject eliminated in their first four moves; the more effective subjects eliminate more objects. Ault did not specify a "correct" object prior to the start of the game, but instead attempted to answer "no" to the first four questions and then bring the game to an end as rapidly as possible. It is interesting to note that if Ault had answered "no" whenever a subject classified less than half of the candidates and if she had answered "yes" whenever they classified over half of the remaining objects, then her measure of 20-questions effectiveness and the E scores would yield (roughly) the same rank ordering.
Secondly, the findings of a number of studies (e.g., Mosher-Hornsby, 1966; Van Horn-Bartz, 1968 and Eimas, 1970) suggest that an effective or optimal strategy for the 20-questions task is simply beyond the capacity of subjects under the 11-12 year old age group. Moreover, Flavell (1963, p. 265) points out that thinking in terms of formal logical relations normally begins to appear at about 11-12 years of age. The participants of the Denney (1973) and Ault (1973) studies were under the 11-12 year old age level. Thus, it appears that even though the reflective subjects of these studies displayed a preference for the use of more sophisticated questions, perhaps (at their age level) they can not concoct a truly effective or optimal line of questioning; hence, the reflectives preference for the basics of a more effective strategy goes for naught.

McKinney (1973), on the other hand, reports that reflective second graders were more effective information gatherers (E scores) than their impulsive peers. However, McKinney used a $2^6$ based grid. Hence, almost all constraint seeking (CS) questions conform to the optimal "half by half. . ." principle and, perhaps, the more effective performance recorded by McKinney's reflectives is a mere consequence of a higher percentage of constraint seeking (CS) questions. Thus, due to the $2^6$ type of array used by McKinney, it is highly likely that his results simply mirror those of Denney (1973) and Ault (1973).

The purpose of the current investigation is: the differential effect of conceptual tempo on the effective use of constraint seeking (CS) questions as maximally efficient information gathering tools. In
other words: Do the participants from particular conceptual tempo groups exhibit a differential preference for hierarchically organizing their line of questioning in accord with the maximally effective "half by half. . ." principle.

The availability of constraint seeking questions is orthogonal and confounding to the purpose of the present investigation. Indeed, if a subject should falter in their implementation of the "half by half. . ." principle, is this state of affairs due to a failure to grasp and actively pursue the benefits of the said principle? Or, is it due--in part--to a lack of necessary constraint seeking (CS) questions? In short, it is difficult to study effectiveness of information gathering prior to the development of a consistently applied constraint seeking (CS) strategy.

Mosher-Hornsby (1966) and Eimas (1970) report that by the eleventh to twelfth year 20-questions respondents begin to rely almost entirely on a constraint seeking strategy. Hence, the lack of constraint seeking (CS) questions dilemma can be greatly alleviated by using subjects from the 11-12 or over age level. Moreover, this dilemma can be further alleviated by (1) the use of stimulus objects with rather simple and highly salient features or attributes and (2) one can heighten the "saliency effect" by reviewing the various levels of each attribute with each participant. In sum, the present investigation is unique and different from the studies of Denney (1973), Ault (1973) and McKinney (1973) in that every attempt has been made to focus on effectiveness of information gathering, while, simultaneously, alleviating the specific hypothesis seeking (HS) versus constraint seeking (CS) issue.
Particulars of 20-Question Arrays

The color-form arrays used in a pilot study appear to fit the aforementioned grid requirements. The Eimas (1970) color-form array served as a prototype for the $2^4$ proportioned color-form grid of the pilot study and hence this array shall be called E'CF-16. Each of the sixteen stimulus objects of E'CF-16 has four attributes fixed at exactly one of two possible levels. The attributes are color (red-blue), shape (square-triangle), letter (B-H) and case of letter (upper case-lower case). The $2^4$ or 16 objects of E'CF-16 were randomly placed into the 4x4 array. The E'CF-16 reprint is on page 34.

Since E'CF-16 was constructed on a $2^4$ basis, one wonders how a participant can fail to behave in accordance with the "half by half..." principle? For instance, if one selects any three of the four attributes and questions on one level for each of the three attributes (e.g., "Is it red?", "Is it a square?" and "Is it a capital letter?") then only two candidates remain. Hence, one has completed stage I with a complete "focusing strategy" (Bruner, et al., 1956; Eimas, 1970) and a perfect $E$ score of 1.

8In the haste of readying the pilot study several of Eimas' original attributes were altered; e.g., Eimas used size of letter where the pilot study used case of letter.

9The 20-questions game was subdivided into two stages. Stage I was primarily characterized by information-gathering questions while Stage II denoted the questioning when only two or three objects remained as candidates for the correct choice. Stage II was characterized by questions designed, not so much for extracting information per se, but for the purpose of guessing the designated "correct" object from the two or three remaining candidates. Roughly, Stage I questions should be constraint seeking (CS) questions while Stage II questions should be specific hypothesis seeking (HS) questions. Thus, the $E$ scores were calculated only for Stage I; the information-gathering stage. Ault (1973) recognized the need for dividing the 20-questions game into two stages when she subdivided the specific hypothesis
Eimas (1970) gives an $E = .85$ for eleven year olds. Although it is not clear if this $E$ is a strictly Stage I mean or not, it illustrates that 11-12 year olds apparently falter somewhere at least some of the time. A pilot study (Spring, 1975) revealed Stage I errors of three types:

1. Redundancy errors; e.g., S asks "Is it red?" E responds "No" and then S asks "Is it blue?"

2. Conjunctive errors; e.g., S asks "Is it a capital B?"
   Thus, the letter and case classes are intersected and E's questioning becomes too specific too soon.

3. Row-column or "rectangular coordinate" error; e.g., S initially categories a single row or column. Note: To initially classify two rows (or columns) and then eliminate one of the two remaining rows (columns) and so on is, indeed, a "complete focusing" or optimal strategy.

Also, there is the potential for a disjunctive error (e.g., "Is is red or blue?"); however, this type of error is a bit deceptive. For instance, a first person may look at one's dog and ask, "What sex is it?" whereas a second might ask, "Is it male or female?" It appears that both people desire the same information and moreover the second person would think it queer if one responded "yes." In other words, the respondents who ask "Is it red or blue?" may not have intended their seeking (HS) questions into two categories; uncompelled HS questions and compelled HS questions. The uncompelled HS questions were Stage I HS questions while the compelled HS questions were Stage II HS questions. Thus, HS-CS questions are not "good" or "bad" per se, but are appropriate or inappropriate depending upon the stage in which they occurred.
question as the union of classes but merely a means of asking "what color is it?" If the respondent's intent fits the latter case then they have not committed a disjunctive error, but ask a nonpermissible (by the rules of the 20-questions game questions must have a yes or no form) question. On the pilot study only one disjunctive error occurred.

The E'CF-16 array, due to the saliency and proportioning of the stimulus objects, provides a number of readily available questions which, in a sense, incidentally cue and guide the respondent to behaving in accordance with the maximally effective "half by half..." principle. The E'CF-16 array does, however, contain a few pitfalls and hence, when attempting this variant of the 20-questions game caution is required. Moreover, and most importantly, the 20-questions game is a task in which a participant can a priori determine "good" moves from "bad" moves and therefore it is possible for the "thinking" player (i.e., the player who considers the validity and outcomes of his actions prior to implementation) to avoid the aforementioned

10 Increasing saliency as used here denotes ready availability of categories that can be used to describe and partition the stimulus array. Eimas (1970) defined increasing saliency as the increasing availability of categories that can be used to describe and partition the stimuli into two equal subsets. Hence, Eimas incorporates both "availability" and "proportionality" in his notion of saliency whereas this writer desires to separate the notions of "readily available" and "proportionality"; for one can separately alter these variables.
pitfalls. Thus, one would expect the E'CF-16 array to favor the reflective subjects. The direction of Ault's (1973) findings suggest that the ascending rank order of the tempo group Stage I $E$ scores would be: impulsives, slow-inaccurates, fast-accurates, and reflectives.

Eimas (1970) demonstrates that 11-12 year-olds (most age groups for that matter) can post rather high $E$ scores on ECT-16 while recording substantially lower $E$ scores on other 16-item 4 by 4 arrays. Moreover, the ECF-16 grid contains highly salient stimulus objects while, comparatively speaking, the other arrays do not. Thus, there is a tendency to claim that high saliency of stimulus objects leads to more effective questioning. This point, however, requires further investigation.

Eimas (1970) points out and his data affirm that "...the ease with which stimuli may be categorized (i.e., partitioned) should be related to the availability of the categories..." The point of conflict arises when one always expects the availability of categories (i.e., saliency in the sense of readily available attributes for categorizing) to elicit the same degree of positive influence on informational efficiency. In short, if the highly salient stimulus objects are proportioned in accord with the "half by half by..." principle then, indeed, high saliency promotes a more effective 20-questions strategy. But, on the other hand, if the highly salient stimulus objects are proportioned in opposition to the "half by half by..." principle then the high saliency condition may elicit a less than optimal 20-questions performance.
In light of the above, a second 16-item color-form array (Niswonger color-form-16 or NCF-16) will be used. NCF-16 consists of stimulus objects which have three attributes (color, shape and letter) and hence E'CF-16 and NCF-16 are somewhat similar. The contrast in these two arrays is provided by the fact that E'CF-16 is proportioned on a $2^4$ basis while NCF-16 is not. The reprint of NCF-16 is on page 39.

NCF-16 consists of 4 levels of color (brown, green, yellow, and red), 3 levels of shape (square, circle and triangle) and 4 levels of letters (A, B, C and D). Findings from the pilot study indicate that about 5/6 of the initial choices go to color or shape with color holding the edge. Thus, in regard to dimensional salience (i.e., the frequency with which particular dimensions or attributes are selected as the basis of categorization) letter is the least salient.

The NCF-16 grid, on the expectation of the above color-shape dominance, was designed so that both color and shape categories allow simple initial questions (i.e., initial questions which do not require disjunctive combinations) which closely approximate the "half by half by..." principle. However, only the letter category provides a simple opening question which exactly fits the maximally effective principle. Initially, all three dimensions contain some relatively ineffective Cs questions and, throughout the game one usually incurs a choice over the relative effectiveness of various Cs questions.

Moreover, the letter category provides the most potential for sequencing a complete focusing strategy from simple Cs questions. The color and shape categories do contain, however, opportunity for the development of near optimal $\bar{E}$ scores. Furthermore, if one does not restrict one's self to simple Cs questions, they can employ disjunctive
combinations of simple color and shape questions to construct a complete focusing strategy. Also, an optimal or complete focusing strategy can be constructed by appropriate spatial partitioning of the array.

E'CF-16 was constructed on the $2^4$ basis while NCF-16 was structured away from the $2^4$ proportioning by placing heightened potential for optimal solution on the attribute of least dimensional saliency. Thus one would expect the Stage I $E$ scores for E'CF-16 to be substantially higher than the corresponding $E$ scores for NCF-16. The results of pilot study-I support the above contention. However, in pilot-I the order of presentation always proceeded from NCF-16 to E'CF-16 and hence these findings could be confounded by an order effect.

The difference in proportioning between NCF-16 and ECF-16 appears to make behavior in accordance to the "half by half. . ." principle a rather simple matter on E'CF-16 while the emergence of such behavior on NCF-16 seems to require more concerted effort on the part of the respondent. However, one would predict that the reflective respondents NCF-16 $E$ scores will be significantly higher than those of their tempo counterparts. Again, the direction of Ault's (1973) findings suggest that the ascending rank order of the tempo group Stage I $E$ scores would be: impulsives, slow-inaccurates, fast-accurates and reflectives.

IQ-Conceptual Tempo - 20 Questions Procedure

While some investigators report (Kagan, et al., 1964) a significant but mild negative correlation between IQ scores and MFF errors and find no significant association between MFF response latency and IQ scores. Other investigators have reported contradictory results.
Denney (1973) reports no substantial relationship between Lorge Thorndike IQ scores and either of the MFF variables. Eska and Black (1971) however, report significant correlations between Otis IQ scores and both of the MFF variables. In short, the conceptual tempo-IQ relationship appears to be dependent upon the particular IQ test.

Eimas (1970) and Denney (1973) report a significant positive correlation between effective 20-questions performance and IQ. Thus, IQ scores on the California Test of Mental Maturity (CTMM) were obtained from school personnel. Moreover, if the correlational results reveal IQ as a potential confounding variable, then this matter will be treated post hoc.

Statement of the Hypotheses

The specific hypotheses to be tested were:

1. To test the null hypothesis $H_0$: the E'CF-16 Stage I $\bar{E}$ means for the four conceptual tempo groups will not significantly ($p < .05$) differ; against the ordered alternative $H_1$: the E'CF-16 Stage I $\bar{E}$ means for the four conceptual tempo groups will be in the ascending order—impulsives, slow-inaccurates, fast-inaccurates, and reflectives—with at least one significant difference.

2. To test the null hypothesis $H_0$: the NCF-16 Stage I $\bar{E}$ means for the four conceptual tempo groups will not significantly ($p < .05$) differ; against the ordered alternative $H_1$: the NCF-16 Stage I $\bar{E}$ means for the four conceptual tempo groups will be in the ascending order—impulsives, slow-inaccurates, fast-accurates, and
reflectives with at least one significant difference.

(3) To test the null hypothesis $H_0$: the E'CF-16 Stage I $\bar{E}$ means for the two orders of presentation (i.e., E'CF-16 first versus E'CF-16 last) will not significantly ($p \leq 0.05$) differ.

(4) To test the null hypothesis $H_0$: the NCF-16 Stage I $\bar{E}$ means for the two orders of presentation (i.e., NCF-16 first versus NCF-16 last) will not significantly ($p \leq 0.05$) differ.

(5) To test the null hypothesis $H_0$: the Stage I $\bar{E}$ means for two 20-questions arrays (E'CF-16 and NCF-16) will not significantly ($p \leq 0.05$) differ; against the ordered alternative $H_1$: the Stage I $\bar{E}$ means for the two 20-questions arrays will be in the ascending order NCF-16, and E'CF-16.

Each of the first four hypotheses will be tested twice--once for the extreme tempo groups and once for the ordinary tempo groups. It is expected that the predicted ordering of hypotheses (1) and (2) will be more prevalent for the extreme groups. Since the ordinary tempo groups were more representative of the population, hypothesis (5) will be tested only for the ordinary tempo groups.
CHAPTER III

PROCEDURE

Introduction

A description of the procedures and considerations for gathering the data required for statistically testing the aforementioned hypotheses is given in this chapter.

In Chapters I and II findings from the pilot studies have been integrated into the "problem development" whenever appropriate. This trend will be continued and pilot results pertinent to procedural matters will be elucidated whenever applicable.

This chapter is organized under the following 3 main sections:
(1) Description of Pupil Sample
(2) Data Collection
(3) Analysis of Data

Description of Pupil Sample

Sample Selection

The participants in this investigation were sixth grade boys from four elementary schools, located in a middle-to-high-income suburb of Columbus, Ohio. The choice of elementary schools was made on the basis of administrative concerns; e.g., feasibility in terms of how heavily the various schools were already committed to outside
projects. The selection of participating schools was based on the underlying assumption that in this relatively homogeneous suburb, the choice of elementary schools would not make much difference.

Initially, three elementary schools were chosen and permission requests were sent (via the children) to parents of all sixth grade boys. The permission slip return rate was slightly in excess of 80 per cent. However, this rate of return did not produce the number of phase-I participants deemed necessary for acquiring equal cell sizes across tempo groups. Thus, a fourth elementary school was included. In all, the four elementary schools yielded a phase-I sample of 140 sixth grade boys.

Data Collection

Phase-I

In phase-I the MFF test (Adult-Adolescent version) was individually administered to the 140 participants (See pages 112 and 113 for sample MFF item). It was estimated that the individual administration of this test would require about 20-30 minutes per child. In order to reduce the phase-I allotment of time, the feasibility of using two Es was considered.

Kagan, et al., (1964) reports that the MFF test results appear to be highly stable and not easily influenced by the contrasting "personalities" of various test administrators. In an attempt to check the stability of the MFF instrument across the two Es, a second pilot study was initiated the week prior to phase-I data collection.
The second pilot study involved 12 randomly selected sixth grade boys from a fifth elementary school. Half of the 12 participants were randomly assigned to one of two test rooms with the remaining half assigned to the other room.

The second pilot involved two days of testing. On the first day one of the two phase-I Es administered the MFF test to the participants of test room I while the other E administered the MFF test to the participants of test room II. On the second day of testing the two Es switched test rooms but the participants reported to the same test rooms as before; thus each subject was tested by each of the two Es.

Since only two days separated the two test periods and since the data revealed a marked drop in both MFF latency and errors on test two, there was little doubt that the results of the second pilot were confounded by a "learning effect." However, the correlation between test one latency and test two latency was .77 while errors correlated at .87 for the two MFF administrations. Thus, it appears that MFF response time and errors are rather stable across the participants, the two Es and the "learning effect." Since these findings concur with Kagan's results (1964), the investigator justified the use of the two Es for phase-I data collection.

The two Es administered the MFF test in each of the four elementary schools. The test rooms varied from vacant classrooms and offices to conference rooms. In some cases the rooms had many windows and overlooked a pleasant view while in other cases the rooms had no windows. However, in all cases the facilities were such that E could
individually administer the MFF test in a private and quiet surrounding.

The MFF test procedure was as follows:

"I am going to show you a picture of a familiar item and then some pictures that look like it. You will have to point to the picture on this bottom page (E points) that is just like the one on this top page (E points). Let's do some for practice." E shows the two practice items and S selects the item that (he thinks) matches. If S was correct, E indicated this to him. If S was wrong, E said "No, that is not the right one, try again. Find the one that is just like this one (E points)."

At the conclusion of the practice items E said "Now we are going to do some more. This time I will be recording information while you match the figures, but just go about the game in your natural manner. These new items are a bit harder. You will see a picture on top and eight pictures on the bottom. Find the one that is just like the one on top and point to it."

E turned the page and said "Start" just as he placed the top page against the backrest. E used a stop watch (kept out of S's view) to record latency of initial response (to nearest half second) for each of the 12 items. E also recorded total number of errors and the order in which they occurred. E continued to code responses (not times) until S made a maximum of eight errors or was correct.

Two MFF procedural matters deserve comment. First, participants would sometimes ask, "Why was my choice wrong?" How does E handle this situation?

Ault (1973) used the technique of "telling a subject why their choice failed to match" as a means of maintaining task interest. She further cited findings (from a pilot study) which indicate that E's verbalization of "what's wrong" is not a sufficient condition to produce a training effect. On these grounds and in order to maintain a friendly rapport, the two Es handled "what's wrong" solicitations by

1Some of the subjects made errors on the practice items, but after a few tries they found the correct match.
replying (in a matter of fact manner), "Such and such; O.K. try again please."

Secondly, although the two Es attempted to keep the stop watch from view, some respondents ask, "Are you timing us?" The two Es handled these requests by answering, "Yes, remember, I said I would be recording information while you worked--but to go about the task in a manner that seems natural." The subjects who displayed an interest in what E was doing did not seem to alter markedly their response style--either prior or after questioning E.

Two participants rather drastically altered their response styles. In both cases, the respondents lengthened their response times and reduced errors. It is interesting to note that in both cases, these respondents verbalized concern over the "poor" quality of their performance just prior to altering their response style. In regard to tempo group, both participants were categorized as slow-inaccurates.

Since the investigator planned to administer the 20-questions tasks, he did not participate in the phase-I data collection. At the completion of phase-I the MFF data sheets were taken to The Ohio State University computer center and turned over to an associate.

Each participant's MFF data sheet contained their subject identification number, mean initial latency of response for 12 MFF items and total errors over all 12 MFF items. The associate ran a scatter plot and calculated the correlation between the two MFF variables.²

²The computer program was Bio Medical Data 02D (BMD 02D) Correlations. Health Sciences Computing Facility, University of California at Los Angeles (UCLA).
The two MFF variables (response latency and total errors) correlated at \(-.58\). The scatter plot revealed that the MFF data lacked linearity.\(^3\) (See Scatter Plot I, page 49).

If one's data does not fit the prescribed linear model, a logarithmic transformation is sometimes used as a remedy (Denney, 1973). Hence, the above program was rerun with a logarithmic transformation on the MFF latency variable and the resulting scatter plot (See Scatter Plot II, page 50) and a correlation between MFF variables of \(-.66\) were obtained.

The conceptual tempo groups were determined by sketching medians on Scatter Plot I (See Scatter Plot III, page 51).\(^4\) Thus, each quadrant corresponds to a tempo group. Since errors were plotted on the vertical axis and latency on the horizontal, it follows that the upper right hand quadrant represents the slow-inaccurates, the upper left hand quadrant represents the fast-inaccurates (impulsives), the lower left hand quadrant represents the fast-accurates and the lower right hand quadrant represents the slow-accurates (reflectives).

\(^3\)The nonlinearity of this curve can be explained by the previously developed (Chapter 2) notion of strategy differences across tempo groups. Consider three types of MFF responders: (1) the person who selects the first figure that appears to match—they respond very rapidly and make many errors; (2) the person who uses a modified search and elimination procedure—they respond somewhat quickly and make some errors; (3) the person who uses a very thorough search and elimination procedure—they respond slowly and make few errors. The (2) type person when compared to the (1) type gains a marked decrease in errors with the expenditure of a bit more time whereas the (3) type person compared to the (2) type can only marginally decrease errors with their added increments of time. In short, the type (3) person is up against a "ceiling effect" or, alternatively, the "law of diminishing returns."

\(^4\)The logarithmic transformation did not alter tempo group classification.
SCATTER PLOT 1. MFF: Latency x Errors.
SCATTER PLOT 2. MFP: Log(Latency) x Errors.
SCATTER PLOT 3. MFF: Latency x Errors, with Median Splits (Reflectives-Impulsives Defined by Strict Inequalities.)
In line with the strong negative correlation between the two MFF variables, one finds the above split median analysis categorized the 140 phase-I participants as follows: 19 slow-inaccurates, 52 impulsives, 19 fast-accurates and 50 reflectives.

Since equal cell sizes has been the goal, it appears that one should randomly select 18 participants from each of the four tempo groups and then randomly assign half of each tempo group to one of the two orders of presentation and the remaining half to the other. However, as one glances at the split-median scatter plot (See scatter plot III, page 51) one notices that the fast-accurates, and the slow-inaccurates tend to cling to the axes of their respective quadrants--much like a rectangular hyperbola--and one wonders if the fast-accurates and slow-inaccurates are borderline reflectives and impulsives that strayed across their respective medians. At any rate, there are reflectives who differ only slightly from fast-accurates, reflectives who differ only slightly from slow inaccurates and reflectives who differ only slightly from fast-accurates, slow-inaccurates and impulsives as well as reflectives who are extremely reflective. Likewise, a similar situation holds for the impulsives. Note: In using four tempo groups, one has increased the number of border lines and hence one has opened more possibilities for reflectives who are near-fast-accurates and so on.

Moreover, if strategy differences (Chapter 2) underlie the various tempo groups then one would expect the preferences for these differential strategies to be more prevalent in the extremes. Hence, if one expects to find E score differences across conceptual tempo groups, it would appear that these differences would be more prominent in the extreme groups.
The above concerns provided support for the decision to test for differences across the extreme groups as well as ordinary groups. The extreme groups were formed by cutting the scatter plot axes at the first, second and third quartile points rather than relying solely on the second quartile or split median analysis (See scatter plot IV; page 54). Thus the extreme reflectives were persons scoring below the 25 percentile on MFF errors and above the 75 percentile on MFF latency whereas the extreme impulsives were persons scoring above the 75 percentile on MFF errors and below the 25 percentile on MFF latency.

Scatter plot IV shows, there were no extreme fast-accurates or slow-inaccurates, the extreme groups classification produced 19 extreme impulsives and 20 extreme reflectives. Thus, the extreme reflectives and impulsives consisted of about 40 per cent of their respective groups.

Again, equal cell sizes were maintained by randomly selecting 18 subjects from each of the extreme groups and then randomly assigning half of each group to one of two orders of presentation and the remaining half to the other. Now that the experimental groups (4 ordinary groups and 2 extreme groups) had been assigned an order of presentation, the random assignment to an order of presentation was extended to the remaining phase-I participants. Again, this was set up so that half of the 50 reflectives and half of the 52 impulsives received one order of presentation and the remaining half received the other. The one remaining fast-accurate was randomly assigned an order of presentation and the remaining slow-inaccurate was assigned the opposite order of presentation.
SCATTER PLOT 4. MFF: Latency x Errors, with Median and 1st and 3rd Quartile Splits (Reflectives-Impulsives Defined by Strict Inequalities.)
The randomization processes were carried out by an associate. The associate preassigned subject identification numbers and order of presentation codes to the phase-II data sheets. Thus the investigator, who was the phase-II E, had no knowledge of the participant's response style.

**Phase II**

Phase II entails the individual administration of the 20-questions procedure to each of the phase-I participants. In phase-II each participant played two twenty questions games on each of the two arrays (E'CF-16 and NCF-16) described in Chapter 2.

On the NCF-16 array the object chosen as "correct" for game one was the brown, triangle with the letter A and the "correct" object for game two was the yellow, circle with the letter D. Likewise, for the ECF-16 array the object chosen as 'correct' for game one was the red-square with a capital B while the "correct" object for game two was the blue, triangle with the lower case b. The designated "correct" objects were chosen so as to prolong the games and thereby allow an adequate number of responses for the calculation of each subject's Ės. Most subject's Ės were calculated over 6-8 E₁ scores; however, a few subjects went as low as 4E₁ scores and some subjects went slightly over 8 E₁ scores.

The specific directions for the 20-questions procedure were:

'We are going to play a game in which I will be thinking of one of the sixteen objects in this 4 by 4 grid (remove the cover to expose first row and column) and your job will be to point out the object that I have in mind. In attempting this task you may ask questions that I can answer with a 'yes' or 'no.' The goal is to point out the object I'm thinking of in
as few questions as possible; the pointing out move will count as a question. When the game starts I will remove the cover, but now I want to give you a chance to become familiar with the features of the exposed objects."

E named and pointed out all levels of all attributes; starting with color, shape, letter and the case. E made it clear that he would be thinking of the whole object—not just the letter part and so on. E again made it clear that the object he had in mind could be any of the 16 objects.

"Let's look at some examples. If a player immediately points to an object and asks "Is this it?'', and provided they are correct, they have completed the game in one question. Now, if a player asks a first question and I answer 'yes' or 'no' and then they point and say "Is this it?'', provided they are correct, they have completed the game in two questions.

We'll review the rules and then start. The pointing out move counts as a question. If a pointing out question "misses" then the player continues and the "miss" counts as a question. All games end when you point out the designated object or if you exceed 12 questions. 12 questions is simply a maximum. Try to do the game in as few questions as possible. As you can see from the examples, "guessing" is permitted; but you be the judge as to if, and when, guessing is to your advantage."

E informs S, "for purposes of clarification, I'm going to record your questions on paper and tape." E starts the tape recorder, removes the cover and says "start."

At the conclusion of game one on the first array E (pauses) says "Let's do this again, I'm thinking of something--start."

At the conclusion of game two on the first array E (pauses) remarks, "Now we are going to play two more games. The rules and goals will be the same as before, but this time I have a slightly different grid." As E made this latter remark he exposed the first row and column of the second grid and then proceeded to go over the various levels of the attributes as before. The procedure for grid two was the same as grid one with the exception of detailed replication of rules and examples.

At the conclusion of the final game the participants were asked if they had a rule or system for playing these games.
At the conclusion of phase-II data collection, each principal gave the investigator a listing of available IQs (CTMM). The list was constructed so that IQs matched the appropriate subject identification number. In all both phases of data collection went well. The investigator had been concerned over the possibility of collaboration among subjects, but the data showed no collaboration effect. The only hitches were that in phase-II, the investigator reversed the order of presentation for two of the reflective subjects, and also one participant was absent from school during the phase-II data collection and hence was dropped from the study. Since none of these subjects were included in the "experimental" groups these errors were of no serious consequence.

Analysis of Data

Data Processing

The investigator took each participant's phase-II data sheet and determined the probability of a "yes" response for each stage I question for both games on both grids (NCF-16 and ECF-16). The investigator then made a listing of each participants stage I probabilities. This listing of probabilities was key-punched on computer cards so that the appropriate $E$ could be calculated by a computer program.

The investigator reviewed the phase-II tapes and recorded the latency of response (nearest half second) on each 20-questions question for the 139 phase-II participants. The latency for the first question
of a game was taken as the time lapse between E's "start" and S identifying a particular category or set of objects. In other words if the subject's initial question was "Is it red?" then the investigator started the stop watch up on hearing "start" and stopped the watch upon hearing "red." Subjects would sometimes ask "Is it ah-ah red?" and one could never be certain if the "ah-ah" (and the "Is it" for that matter) was a part of their thinking time or not. On the non-game opening questions the response time was taken as the time lag between E's "yes" or "no" answer and S identifying the next particular category.

In all there were two 20-questions latency variables for each of the two grids. In the 20-questions instructions, the subjects were never told the proportions of the various levels of the attributes. Hence, to employ the "half by half by. . ." principle most effectively, one would likely need to count the number of browns, greens and so on. Thus, differences in latency of initial response could be revealing. However, for a given grid the only truly virgin initial response occurs on the first question of the first game and hence this was one of the 20-questions latency measures. This variable was denoted by 20-Q-latency-I.

Participants who play the 20-questions game in accordance to the optimal "half by half by. . ." strategy require three questions to complete stage-I. On the other hand, participants who do not play in accordance with the optimal strategy often required more than (on occasion less than) two non-initial stage-I questions. However, it appears that response time on these extra stage-I questions is more apt
to be contaminated with "being baffled" time lag rather than "planning" or "thinking about what one's doing" latency. Thus, to acquire a measure of stage-I (non-initial question) per question expenditure of time, it seemed natural to take the average latency of a subject's second and third stage-I questions. This measure of 20-questions latency was denoted by 20-Q-latency-II.

The purpose of this investigation was the differential effect of conceptual tempo on the effectiveness of 11-12 year olds information gathering strategies across several rather similar 20-questions arrays. Thus a subject's \( \bar{E} \) for each grid was the stage-I grand mean for the given grid; i.e., the mean obtained from pooling the subject's stage-I \( E_i \)'s for the given grid. Likewise, the per grid 20-Q-latency-II variable was the grand mean for the given grid. The per grid 20-Q-latency-I variable was the latency of initial response (first game) for the given grid.

In all, nine experimental variables (\( \bar{E} \)-E'CF-16, \( \bar{E} \)-NCF-16, E'CF-16-latency-I, E'CF-16-latency-II, NCF-16-latency-I, NCF-16-latency-II, IQ(GTMM), MFF Errors, MFF latency) along with subject identification numbers and classification codes (tempo group, order of presentation and grid) were key punched onto computer cards.

**Data Analysis Procedures**

\( \bar{E} \)-E'CF-16 was used as the criterion variable in two 4(group) by 2 (order of presentation) univariate analysis of variance procedures. One for ordinary tempo groups and the other for extreme tempo groups. Likewise, \( \bar{E} \)-NCF-16 was used as the criterion variable in two analogous
procedures. After discerning the extent of the order effect, $E-E'CF-16$ and $E-NCF-16$ were appropriately collapsed into a single informational effectiveness variable $\bar{E}$, and used as the criterion variable in the $4 \times 2$ (ordinary group) by $2 \times 2$ (grid) univariate analysis of variance procedure. The ANOVA procedures were carried out by use of the computer program Statistical Package for the Social Sciences (SPSSG031).

The aforementioned nine experimental variables were subjected to a correlational analysis. These correlations were carried out by the computer program "Spearman Rank Corr" Statistical Package for the Social Sciences (SPSSG031).
CHAPTER IV

RESULTS

Introduction

The results of this investigation are presented in the order of the listing in the Data Analysis Procedures section of Chapter III. The results are organized under the following six main sections:

(1) $\bar{E}$-E'CF-16; the results of the four tempo by two orders of presentation univariate analysis of variance on the criterion variable $\bar{E}$-ECF-16.

(2) $\bar{E}$-NCF-16; the results of the four tempo by two orders of presentation univariate analysis of variance on the criterion variable $\bar{E}$-NCF-16.

(3) NCF-16 versus E'CF-16; the results of the four tempo by two grid univariate analysis of variance on the criterion variable $\bar{E}$.

(4) Conclusions; concerning the stated hypotheses.

(5) Correlational Analysis

(6) Summary

The MFF results were presented in the phase-I part of the data collection section of Chapter III and hence will not be duplicated in Chapter IV.

$\bar{E}$-ECF-16

Extreme Groups

Analysis of the mean informational effectiveness scores (amount of information--measured in bits--per stage I question) of the eight
cells formed by crossing the four tempo groups with the two orders of presentation revealed that, regardless of order of presentation, the extreme reflectives were somewhat more effective than the slow-inaccurates, the slow-inaccurates and the fast-accurates were practically equally effective while the extreme impulsives were slightly less effective than the fast-accurates (See Table 1 and Graph 1). Moreover, the extreme reflectives, the slow-inaccurates and the fast-accurates evidenced marginally lower mean effectiveness scores when E'CF-16 was preceded by NCF-16 whereas the extreme impulsives displayed slightly higher mean effectiveness scores when E'CF-16 was presented second.

**TABLE 1**

INFORMATIONAL EFFECTIVENESS-EXTREME GROUPS
CELL MEANS, STANDARD DEVIATIONS AND SKEWNESS FOR E-E'CF-16

<table>
<thead>
<tr>
<th></th>
<th>Extreme</th>
<th></th>
<th>Extreme</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Slows</td>
<td>Impulsives</td>
<td>Quicks</td>
</tr>
<tr>
<td></td>
<td>Std.</td>
<td>Std.</td>
<td>Std.</td>
</tr>
<tr>
<td></td>
<td>Mean Dev.</td>
<td>Skew</td>
<td>Mean Dev.</td>
</tr>
<tr>
<td>ECF-16</td>
<td>0.95</td>
<td>0.05</td>
<td>-0.57</td>
</tr>
<tr>
<td>First</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ECF-16</td>
<td>0.92</td>
<td>0.08</td>
<td>-1.00</td>
</tr>
<tr>
<td>Second</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The E'CF-16 array had the convenient $2^4$ proportioning and overall the E-E'CF-16 scores were near optimal—especially the extreme reflective subjects. These results presented some special problems. First, the cell distributions lacked normality (See Skewness-Table 1). Second, there was a lack of cellwise homogeneity of variance—particularly the extreme impulsives compared with the extreme reflectives.
GRAPH 1. Extreme Groups $\bar{E}-E'\text{CF-16}$ 4 Tempo Groups x 2 Orders of Presentation.
when E'CF-16 was presented first. However, the cell sizes were equal and as Brown and Forsythe (1974) point out (with the variances at hand) this prevents any serious consequences due to the lack of homogeneity of variance.

Now, back to the normality issue. Although analysis of variance tests are generally viewed as rather "robust" against non-normality (e.g., Guilford, 1965), under the conditions at hand, it is more appropriate to use a distribution-free nonparametric test. Hettmansperger's (1975) Non-Parametric Inference test for Ordered Alternatives in a Randomized Block Design appears to best fit the current situation. However, Hettmansperger's (1975) test does not provide a ready means for testing an interaction effect; thus, both the analysis of variance procedure and the Hettmansperger test were performed on the data.

First, the Hettmansperger test. Table 2 gives the two-way layout for testing the tempo group main effect with the two orders of presentation as blocking variables. Table 3 contains the sums of

---

1 The nonparametric tests were reported only when the results between nonparametric and parametric differed. In these cases the nonparametric results were considered the more appropriate (due to the normality issue) and they (nonparametric) were reported in lieu of the analysis of variance findings. Only one such case occurred and that involved the extreme group effect on E-E'CF-16; the case where the normality issue was the most severe.

2 In the Hettmansperger test there is little to be gained from equal cell sizes (save, perhaps, some stability) and it appears that a much truer picture would emerge from a ranking of the entire phase-II population. Thus, the 139 phase-II participants were included in the nonparametric tests. Under these conditions, the impulsives and reflectives were split into two groups; extreme impulsives as before and impulsives who were nonextreme (denoted impulsives-nonextreme); likewise for the reflectives. The fast-accurate and slow-inaccurate groups were as before. However, it is important to note that the impulsives-nonextreme and reflectives-nonextreme differ from the
## TABLE 2

**DISPLAY OF E-ECF-16 SCORES AND RANKS WITHIN THE TEMPO GROUPS AND THE RESPECTIVE ORDER OF PRESENTATION BLOCKS. NUMBERS IN PARENTHESES DENOTE RANKS WHILE THE NUMBERS IN BRACKETS INDICATE THE NUMBER OF TIES.**

<table>
<thead>
<tr>
<th>Impulsive Extreme</th>
<th>Impulsive Nonextreme</th>
<th>Slow-Inaccurates</th>
<th>Fast-Accurates</th>
<th>Reflectives Nonextreme</th>
<th>Reflectives Extreme</th>
</tr>
</thead>
<tbody>
<tr>
<td>ECF-16 First</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.0000(56.5)[1]</td>
<td>1.000 (56.5)[6]</td>
<td>1.0 (56.5)[3]</td>
<td>1.0 (56.5)[4]</td>
<td>1.0 (56.5)[7]</td>
<td>1.0 (56.5)[9]</td>
</tr>
<tr>
<td>.9371(22.5)[3]</td>
<td>.9685(36.5)[1]</td>
<td>.9685(36.5)[2]</td>
<td>.9685(36.5)[3]</td>
<td>.9685(36.5)[4]</td>
<td>.9371(22.5)[1]</td>
</tr>
<tr>
<td>.9344(17 )[1]</td>
<td>.9549(28 )[3]</td>
<td>.9371(22.5)[1]</td>
<td>.9614(30.5)[1]</td>
<td>.9614(30.5)[1]</td>
<td>10</td>
</tr>
<tr>
<td>.8958(12 )[1]</td>
<td>.9344(17 )[2]</td>
<td>.9215(14 )[1]</td>
<td>.8582( 9 )[1]</td>
<td>.9371(22.5)[3]</td>
<td></td>
</tr>
<tr>
<td>.8606(10.5)[1]</td>
<td>.9235(15 )[1]</td>
<td>.9088(13 )[1]</td>
<td>.7264( 3 )[1]</td>
<td>.8469( 6 )[1]</td>
<td></td>
</tr>
<tr>
<td>.7500( 4 )[1]</td>
<td>.8571( 7.5)[2]</td>
<td>.8606(10.5)[1]</td>
<td></td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>.5995( 2 )[1]</td>
<td>.3743( 1 )[1]</td>
<td></td>
<td></td>
<td>9</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>9</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>16</td>
<td></td>
</tr>
</tbody>
</table>

| ECF-16 Second     |                      |                  |                |                        |                     |
| 1.0 (59 )[1]      | 1.0 (59 )[4]         | 1.0 (59 )[2]     | 1.0 (59 )[2]    | 1.0 (59 )[4]           | 1.0 (59 )[6]        |
| .9685(45 )[1]     | .9685(45 )[2]        | .9685(45 )[2]    | .9561(38 )[1]   | .9685(45 )[4]          | .9371(34.5)[1]      |
| .9371(34.5)[2]    | .9623(40 )[1]        | .9344(30.5)[1]   | .9371(34.5)[2] | .9614(39 )[1]          | .9227(25 )[1]       |
| .9324(28.5)[1]    | .9343(34.5)[1]       | .9082(23 )[1]    | .9245(27 )[1]   | .9324(28.5)[1]         | .8176( 9.5)[1]      |
| .9235(26 )[1]     | .9344(30.5)[1]       | .8986(21 )[1]    | .8571(15 )[2]   | .9099(24 )[1]          | .7772( 5 )[1]       |
| .9074(22 )[1]     | .8571(15 )[3]        | .8688(19 )[1]    | .8032( 7 )[1]   | .8032(11 )[1]          | 10                  |
| .8710(20 )[1]     | .8469(12 )[1]        | .8606(18 )[1]    |                | 9                      |                     |
| .8302(11 )[1]     | .8176( 9.5)[1]       | .7500( 3.5)[1]   |                | .7029( 1 )[1]          | 13                  |
| .7264( 2 )[1]     | .8032( 7 )[1]        | .7500( 3.5)[1]   |                | 10                     |                     |
|                   |                      |                  |                | 16                     |                     |
the average cell rank for each tempo group.

### Table 3

<table>
<thead>
<tr>
<th></th>
<th>Impulsive</th>
<th>Impulsive</th>
<th>Slow-</th>
<th>Fast-</th>
<th>Reflective</th>
<th>Reflective</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Extreme</td>
<td>Nonextreme</td>
<td>Inaccurate</td>
<td>Accurate</td>
<td>Nonextreme</td>
<td>Extreme</td>
</tr>
<tr>
<td>Sum of Cell Rank Averages per Tempo Group</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Impulsive Extreme</td>
<td>47.1</td>
<td>69.6</td>
<td>65.9</td>
<td>67.2</td>
<td>78.2</td>
<td>95.9</td>
</tr>
</tbody>
</table>

Briefly, according to Hettmansperger (1975) in a (b) blocks by (c) treatments (groups, in the case at hand) experiment:

\[
T = \sum_{c}^{b} \sum_{j=1}^{i} R_{ij}/n_{ij}
\]

where \( n_{ij} \) denotes the number of observations in the \( i, j \) cell (\( j \) reflects the hypothesized ordering; the largest value going to the most favored group) and \( R_{ij} \) denotes the sum of the ranks of the observations in that cell. Furthermore, under the null hypotheses the mean and variance of \( T \) are respectively given by:

\[
\text{ET.} = \frac{C(C + 1)(N + b)}{4}
\]

where \( N \) denotes the total number of observations.

---

previously described ordinary impulsives and ordinary reflectives. The term "ordinary" denotes a random selection of subjects from the entire impulsive or reflective group.
(3)\( \text{Var } T = \sum_{i=1}^{b} \sum_{j=1}^{c} \frac{N_i(N_i+1)j^2}{12} n_{ij} - C^2(C+1)^2(N+b)/48 \)

where \( N_i \) denotes the number of observations in the \( i^{th} \) block.

Since a directional hypothesis is being tested, the test is one sided and with \( \alpha = .05 \) \( H_0 \) is rejected if

\[
\frac{(T - ET)}{(\text{Var } T)^{1/2}} > 1.645
\]

The directional hypothesis was:

- extreme impulsives \( \leq \) impulsives nonextreme \( \leq \) slow-inaccurates
- fast-accurates \( \leq \) reflectives nonextreme \( \leq \) extreme reflectives

(with at least one strict inequality). Thus, \( H_0 \) is rejected since \((1619.2 - 1480.5)/(1583)^{1/2} = 3.46.\)

The prime interest has been in the extreme groups and the two middle groups (i.e., fast-accurates and slow-inaccurates). Thus, Table 4 contains multiple comparisons only for these groups. Moreover, if the fast-accurate vs. slow-inaccurate comparison is ignored then the overall \( p < .05 \) level can be maintained by conducting the five individual pairwise comparisons at the \( p < .01 \) level.

Again, via Hettmansperger\(^3\) (1975):

\[
D_{jk} = \sum_{i=1}^{b} \left( \frac{R_{ik}/n_{ik}}{R_{ij}/n_{ij}} \right)
\]

for \( j < k \). Then under the null hypothesis: \( ED_{jk} = 0 \) and

\(^3\)Simply put, \( D_{46} \) indicates the difference between fast-accurates and extreme reflectives in regard to sums (across blocks) of cell rank averages (Table 3); i.e., \( D_{46} = 28.7 \).
\[ \text{Var } D_{jk} = \Sigma b \frac{N_i (N_i + 1) (n_{ik}^{-1} + n_{ij}^{-1})}{12}. \]

Declare the effect of treatment K to be significantly greater than the effect of treatment j (j < k) if \( D_{jk} > z \text{ (var } D_{jk} \text{)}^{1/2} \) where z is the normal standard score for the prescribed (pairwise, .01 in this case) \( \alpha \) level.

**TABLE 4**

**MULTIPLE COMPARISONS**

<table>
<thead>
<tr>
<th>Groups</th>
<th>( D_{jk} )</th>
<th>( Z(\text{Var } D_{jk})^{1/2} )</th>
<th>( P \leq .05 )</th>
<th>( P \leq .1 )</th>
<th>( P \leq .05 )</th>
<th>( P \leq .1 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extreme Reflective vs.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fast-Accurates</td>
<td>28.7</td>
<td>30.6</td>
<td>27.2</td>
<td>No</td>
<td>Yes*</td>
<td></td>
</tr>
<tr>
<td>Slow-Inaccurates</td>
<td>30.0</td>
<td>30.3</td>
<td>26.9</td>
<td>No</td>
<td>Yes*</td>
<td></td>
</tr>
<tr>
<td>Extreme Reflective vs.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Extreme Impulsive</td>
<td>48.8</td>
<td>30.3</td>
<td>26.9</td>
<td>Yes*</td>
<td>Yes*</td>
<td></td>
</tr>
<tr>
<td>Fast Accurates vs.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Extreme Impulsive</td>
<td>20.1</td>
<td>30.6</td>
<td>27.2</td>
<td>No</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>Slow-Inaccurates vs.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Extreme Impulsives</td>
<td>18.8</td>
<td>30.6</td>
<td>27.2</td>
<td>No</td>
<td>No</td>
<td></td>
</tr>
</tbody>
</table>

* Denotes the comparisons that were significant

Furthermore, the differences between the extreme reflectives vs. fast-accurates and extreme reflectives vs. slow-inaccurates are both significant at \( P \leq .073 \). However, since the Hettmansperger method does not include a correction for the over estimation of variances due to tied scores, one must consider the .073 level of significance a rather conservative upperbound. Moreover, the high number of ties within the
extreme reflective, fast-accurate and slow-inaccurate groups, suggests that it is quite reasonable to believe that the aforementioned differences fall within the \( P < .05 \) level of significance (See Hollander and Wolfe, 1972, p. 116). The current nonparametric literature contains no such "ties" correction factor for the Hettmansperger test and at present this is the best argument that one can offer.

Thus, on the grounds of the above, one should tentatively hold that on the E'CF-16 grid the extreme reflectives were significantly more effective (stage I) information gatherers than the fast-accurates and slow-inaccurates. In all, on E'CF-16 the extreme-reflectives were significantly more effective stage I information gatherers (i.e., the hierarchical organization of their line of questioning more closely adhered to the optimal "half by half. . ." principle) than the extreme impulsives, slow-inaccurates and fast-accurates.

Ignoring order of presentation and simply counting the number of optimal scores (what Eimas, 1970 and Bruner, 1956 call complete conservative focusing strategies) within each tempo group of Table 2 one finds that the extreme reflectives exhibited a 75% optimal strategy rate, while the extreme impulsives displayed a 10% optimal strategy rate and the remaining tempo groups evidenced a 33% optimal strategy rate. This provides further support for the view that when it comes to avoiding the "pit falls" of E'CF-16, the extreme reflectives were substantially superior to the other tempo groups.
The four tempo by two orders of presentation univariate analysis of variance procedure on the criterion variable $\bar{E}-E'CF-16$ revealed a significant ($p < .01$) main effect exclusively for conceptual tempo, and no significant tempo by order of presentations interaction (See Table 5). In summary, the significant tempo effect is in agreement with the previous discussion $^4$ while the non-significant order of presentations effect concurs with a second nonparametric test. $^5$

**TABLE 5**

TWO WAY ANALYSIS OF VARIANCE FOUR TEMPO X TWO ORDERS OF PRESENTATION $\bar{E}-E'CF-16$-EXTREME GROUPS

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>P less than</th>
</tr>
</thead>
<tbody>
<tr>
<td>Order</td>
<td>1</td>
<td>.005</td>
<td>0.739</td>
<td>.999</td>
</tr>
<tr>
<td>Tempo Group</td>
<td>3</td>
<td>.026</td>
<td>4.136</td>
<td>.010</td>
</tr>
<tr>
<td>Order x Tempo</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group Interaction</td>
<td>3</td>
<td>.003</td>
<td>0.521</td>
<td>.999</td>
</tr>
</tbody>
</table>

$^4$Although the Hettmansperger and analysis of variance tests were in agreement in regard to the results of the conceptual tempo main effect, differences emerge when the multiple comparisons are considered. The Tukey-HsD Multiple Range test revealed significant differences ($p < .05$) exclusively for the extreme impulsives vs. extreme reflectives comparison whereas the Hettmansperger multiple comparisons revealed substantial differences ($p < .073$—conservative upper bound—which was believed to be within the .05 level) which indicated that the extreme reflectives were markedly superior to, not only the extreme impulsives, but the slow-inaccurates and fast-accurates as well. Moreover, the Hettmansperger results concurred with the "optimal strategy rate" findings (See page 69).

$^5$The "order of presentation" hypotheses (hypotheses 3 and 4) were non-ordered hypotheses and hence required a test different from the Hettmansperger test. Here a test developed by Skillings (1975) was used. Roughly, this procedure entailed the computation of a Kruskal-Wallis value for each of the six tempo blocks (now the tempo groups are the blocking variables and the orders of presentation are the treatment variables); then summing the Kruskal-Wallis values and
Ordinary Groups

An examination of the mean informational effectiveness scores of the eight cells of the four tempo by two orders of presentation ANOVA revealed that, regardless of order of presentation, the reflectives were marginally more effective than the slow-inaccurates—although the slow-inaccurates, fast-accurates and impulsives were ranked (descendingly) in the order listed—these latter three groups were practically equally effective (See Table 6 and Graph 2). The means of the reflective groups held constant across the two orders of presentation while the other groups evidenced slightly lower mean effectiveness scores when E'CF-16 was preceded by NCF-16.

TABLE 6

INFORMATIONAL EFFECTIVENESS-ORDINARY GROUPS
CELL MEANS, STANDARD DEVIATIONS AND SKEWNESS
E-E'CF-16

<table>
<thead>
<tr>
<th>Slow-Inaccurates</th>
<th>Impulsives</th>
<th>Fast-Accurates</th>
<th>Reflectives</th>
</tr>
</thead>
<tbody>
<tr>
<td>ECF-16 First</td>
<td>0.95</td>
<td>0.05</td>
<td>-0.57</td>
</tr>
<tr>
<td>ECF-16 Second</td>
<td>0.92</td>
<td>0.08</td>
<td>-1.00</td>
</tr>
</tbody>
</table>

checking a Chi square table for six degrees of freedom.
Note: Since there were two orders of presentation, one degree of freedom was gained on each of the six tempo blocks.
GRAPH 2. Ordinary Groups E-E'CF-16 4 Tempo Groups x 2 Orders of Presentation.
The four tempo by two orders of presentation univariate analysis of variance procedure on the criterion variable E-E'CF-16 revealed no significant \( p < .05 \) main effects or interactions (See Table 7). The results, however, were in the predicted direction and the two way ANOVA revealed a tempo effect at the \( p < .10 \) level of significance. (See Table 7) The nonparametric test results (i.e., tempo group and order of presentation effects) concurred with the findings reported in Table 7. Thus, the nonparametric results were not reported.

**TABLE 7**

**TWO WAY ANALYSIS OF VARIANCE FOUR TEMPO X TWO ORDERS OF PRESENTATION E-E'CF-16-ORDINARY GROUPS**

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>Mean Square</th>
<th>( F )</th>
<th>( P ) less than</th>
</tr>
</thead>
<tbody>
<tr>
<td>Order</td>
<td>1</td>
<td>0.01</td>
<td>2.13</td>
<td>0.146</td>
</tr>
<tr>
<td>Tempo Group</td>
<td>3</td>
<td>0.01</td>
<td>2.27</td>
<td>0.088</td>
</tr>
<tr>
<td>Order x Tempo Group</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interaction</td>
<td>3</td>
<td>0.001</td>
<td>0.213</td>
<td>0.999</td>
</tr>
</tbody>
</table>

For both the reflective and the impulsive groups, the extremes comprised about forty percent of the total group. The randomly selected ordinary reflectives consisted of fifty percent extremes while the randomly selected ordinary impulsive group contained approximately thirty percent extremes. Thus, in regard to extremes,\(^6\)

\(^6\)The Hettmansperger (1975) test showed a tempo group effect with a \( p < .098 \) significance level. Again, the Hettmansperger test lacked the "ties" correction factor; but, the point remains, the inclusion of the nonextremes (reflective-impulsive) waters down the significance of the differences.
the ordinary reflectives had a ten percent surplus while the ordinary impulsives had a ten percent deficit and therefore these randomly selected groups (relatively speaking) adequately represented their respective populations.

**E-NCF-16**

**Extreme Groups**

Analysis of the mean informational effectiveness scores (amount of information—measured in bits—per Stage I question) of the eight cells formed by crossing the four tempo groups with the two orders of presentation revealed that when NCF-16 was presented first the extreme reflectives, fast-accurates, slow-inaccurates and extreme impulsives were ranked (descendingly) in the order listed; however, these groups were practically equally effective. On the other hand, when NCF-16 was presented second the extreme reflectives, slow-inaccurates and extreme impulsives exhibited marginally lower mean effectiveness scores while the fast-accurates displayed marginally higher mean effectiveness scores. The extreme reflectives decrease in informational effectiveness was slightly more severe than the other groups and hence the descending rank order with NCF-16 presented second was: fast-accurates, slow-inaccurates, extreme reflectives and extreme impulsives (See Table 8 and Graph 3).

The four tempo by two orders of presentation univariate analysis of variance procedure on the criterion variable E-NCF-16 revealed no significant \( p < .05 \) main effects or interactions (See Table 9). These results concurred with the nonparametric findings; hence, the nonparametric results were not reported.
TABLE 8
INFORMATIONAL EFFECTIVENESS-EXTREME GROUPS CELL MEANS, STANDARD DEVIATIONS AND SKEWNESS-E-NCF-16

<table>
<thead>
<tr>
<th></th>
<th>Slow-</th>
<th>Extreme</th>
<th>Fast-</th>
<th>Extreme</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Inaccurates</td>
<td>Impulsives</td>
<td>Accurates</td>
<td>Reflectives</td>
</tr>
<tr>
<td>Std. Mean Dev. Skew</td>
<td>Std. Mean Dev. Skew</td>
<td>Std. Mean Dev. Skew</td>
<td></td>
<td></td>
</tr>
<tr>
<td>First</td>
<td>0.88 0.07 -0.69 0.87 0.06 -0.60 0.89 0.07 -0.61 0.90 0.05 -0.53</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NCF-16 Second</td>
<td>0.86 0.06 -0.76 0.84 0.08 -0.66 0.91 0.05 -0.81 0.85 0.09 -0.38</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

TABLE 9
TWO WAY ANALYSIS OF VARIANCE FOUR TEMPO X TWO ORDERS OF PRESENTATION E-NCF-16-EXTREME GROUPS

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>P less than</th>
</tr>
</thead>
<tbody>
<tr>
<td>Order</td>
<td>1</td>
<td>0.009</td>
<td>2.073</td>
<td>0.151</td>
</tr>
<tr>
<td>Tempo Group</td>
<td>3</td>
<td>0.007</td>
<td>1.663</td>
<td>0.183</td>
</tr>
<tr>
<td>Order x Tempo Group</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interaction</td>
<td>3</td>
<td>0.003</td>
<td>0.728</td>
<td>0.999</td>
</tr>
</tbody>
</table>

Ordinary Groups

An examination of the mean informational effectiveness scores of the eight cells formed by crossing the four tempo groups with the two orders of presentation revealed that when NCF-16 was presented first the means of the four tempo groups clustered just below .90. However, when NCF-16 was presented second, the fast-accurates evidenced a slightly higher mean effectiveness score while the other three groups
GRAPH 3. Extreme Groups E-NCF-16 4 Tempo x 2 Orders of Presentation.
evidenced marginally lower mean effectiveness scores (See Table 10 and Graph 4).

TABLE 10

INFORMATIONAL EFFECTIVENESS-ORDINARY GROUPS CELL MEANS, STANDARD DEVIATIONS AND SKEWNESS-Ł-NCF-16

<table>
<thead>
<tr>
<th></th>
<th>Slow- Inaccurates</th>
<th>Extreme Impulsives</th>
<th>Fast- Accurates</th>
<th>Extreme Reflectives</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Std. Mean Dev. Skew</td>
<td>Std. Mean Dev. Skew</td>
<td>Std. Mean Dev. Skew</td>
<td>Std. Mean Dev. Skew</td>
</tr>
<tr>
<td>NCF-16</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>First</td>
<td>0.88 0.07 -0.69 0.88 0.04 0.07 0.89 0.07 -0.61 0.89 0.06 -0.76</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NCF-16</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Second</td>
<td>0.86 0.06 -0.76 0.88 0.09 -1.83 0.91 0.05 -0.81 0.87 0.05 -0.34</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

On the criterion variable Ł-NCF-16 the four tempo by two orders of presentation ANOVA revealed no significant (p < .05) main effects or interactions (See Table 11). These results concurred with the nonparametric findings. Hence the nonparametric results were not reported.

TABLE 11

TWO WAY ANALYSIS OF VARIANCE FOUR TEMPO X TWO ORDERS OF PRESENTATION Ł-NCF-16-ORDINARY GROUPS

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>P less than</th>
</tr>
</thead>
<tbody>
<tr>
<td>Order</td>
<td>1</td>
<td>0.001</td>
<td>0.228</td>
<td>0.999</td>
</tr>
<tr>
<td>Tempo Group</td>
<td>3</td>
<td>0.003</td>
<td>0.795</td>
<td>0.999</td>
</tr>
<tr>
<td>Order x Tempo Group</td>
<td>3</td>
<td>0.002</td>
<td>0.444</td>
<td>0.999</td>
</tr>
</tbody>
</table>
GRAPH 4. Ordinary Groups E-NCF-16 4 Tempo x 2 Orders of Presentation.
**NCF-16 versus E'CF-16**

**Ordinary Groups**

Order of presentation was eliminated from the grid comparisons by restricting the analysis to participants who encountered E'CF-16 first versus the participants who faced NCF-16 first. An examination of the cell means (informational effectiveness—measured in bits of information extracted per stage I question) of the four tempo by two grid ANOVA revealed that the E'CF-16 tempo group cell means ranged from .93 to .97 whereas the NCF-16 tempo group cell means clustered around .88 (See Table 12 and Graph 5).

**TABLE 12**

INFORMATIONAL EFFECTIVENESS (E): E'CF-16 Versus NCF-16
CELL MEANS STANDARD DEVIATIONS AND SKEWNESS
FOR FOUR TEMPO BY TWO GRID ANOVA

<table>
<thead>
<tr>
<th>Slow-</th>
<th>Extreme</th>
<th>Fast-</th>
<th>Extreme</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inaccurates</td>
<td>Impulsives</td>
<td>Accurates</td>
<td>Reflectives</td>
</tr>
<tr>
<td>Std.</td>
<td>Std.</td>
<td>Std.</td>
<td>Std.</td>
</tr>
<tr>
<td>Mean Dev.</td>
<td>Skew</td>
<td>Mean Dev.</td>
<td>Skew</td>
</tr>
<tr>
<td>ECF-16 0.95 0.05 -0.57 0.93 0.05 -0.49 0.94 0.09 -1.55 0.97 0.05 -1.70</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NCF-16 0.88 0.07 -0.69 0.88 0.04 0.07 0.89 0.07 -0.61 0.89 0.06 -0.76</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

On the criterion variable informational effectiveness (E) the four tempo by two grid ANOVA revealed a significant (p<.001) main effect exclusively for grid and no significant (p<.05) interactions (See Table 13). Again, these results concurred with the nonparametric findings. Thus, the nonparametric results were not reported.
GRAPH 5. Ordinary Groups: NCF-16 vs. E'CF-16 4 Tempo Groups x 2 Arrays.
TABLE 13
TWO WAY ANOVA FOUR TEMPO X TWO GRID
E ON ORDINARY GROUPS

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>P less than</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grid</td>
<td>1</td>
<td>0.073</td>
<td>10.482</td>
<td>0.001</td>
</tr>
<tr>
<td>Tempo Group</td>
<td>3</td>
<td>0.002</td>
<td>0.614</td>
<td>0.999</td>
</tr>
<tr>
<td>Grid x Tempo Group</td>
<td>Interaction</td>
<td>3</td>
<td>0.001</td>
<td>0.226</td>
</tr>
</tbody>
</table>

From Table 13 and Graph 5 it is clear that (as predicted) more effective Stage-I information gathering was evidenced on E'CF-16 than NCF-16.

Conclusions: Concerning the Stated Hypotheses

Restatement of Hypotheses

The specific hypotheses were:

(1) To test the null hypothesis $H_0$: the E'CF-16 Stage I $\bar{E}$ means for the four conceptual tempo groups will not significantly ($p < .05$) differ; against the ordered alternative $H_1$: the ECF-16 Stage-I $\bar{E}$ means for the four conceptual tempo groups will be in the ascending order—impulsives, slow-inaccurates, fast-inaccurates, and reflectives—with at least one significant difference.

(2) To test the null hypothesis $H_0$: the NCF-16 Stage I $\bar{E}$ means for the four conceptual tempo groups will not significantly ($p < .05$) differ; against the ordered alternative $H_1$: the NCF-16
Stage I $\bar{E}$ means for the four conceptual tempo groups will be in the ascending order—impulsives, slow-inaccurates, fast-accurates, and reflectives—with at least one significant difference.

(3) To test the null hypothesis $H_0$: the E'CF-16 Stage I $\bar{E}$ means for the two orders of presentation (i.e., E'CF-16 first versus E'CF-16 last) will not significantly ($p < .05$) differ.

(4) To test the null hypothesis $H_0$: the NCF-16 Stage I $\bar{E}$ means for the two orders of presentation (i.e., NCF-16 first versus NCF-16 last) will not significantly ($p < .05$) differ.

(5) To test the null hypothesis $H_0$: the Stage I $\bar{E}$ means for the two 20-questions arrays (E'CF-16 and NCF-16) will not significantly ($p < .05$) differ; against the ordered alternative $H_1$: the Stage I $\bar{E}$ means for the two 20-questions arrays will be in the ascending order NCF-16 and E'CF-16.

Each of the first four hypotheses were tested twice. Once for the extreme tempo groups and once for ordinary tempo groups. Hypothesis (5) was tested only for ordinary tempo groups.

**Conclusions**

The results of section (1) $\bar{E}$-E'CF-16 indicate that on E'CF-16 extreme reflective subjects were significantly ($p < .05$) more effective information gatherers than their extreme impulsive counterparts. The nonparametric multiple comparisons for the tempo effect revealed that the extreme reflectives were significantly ($p < .073$) more effective information gatherers than the slow-inaccurates and the fast-accurates. Since the Hettmansperger (1975) test lacked a "ties"
correction factor the .073 level of significance was viewed as a rather conservative upper bound. Moreover, there was good reason to believe that the 'true' level of significance may be within the .05 level. The optimal strategy rate data added further support to this belief. Thus on these grounds, the original conclusion was distended to read: on E'CF-16 the extreme reflective subjects were substantially more effective information gatherers than their slow-inaccurate, fast-accurate and extreme impulsive counterparts.

Although the trend of reflectives being more effective information gatherers on E'CF-16 than their impulsive counterparts was also exhibited over the ordinary groups, the level of significance (p < .1) was not sufficient to warrant rejection of the null hypothesis. Thus, null hypothesis (1) was rejected only for the extreme groups design.

The results of section (2) E-NCF-16 and section (3) E'CF-16 versus NCF-16 indicate that null hypothesis (5) was the only other null hypothesis that warrants rejection. In the case of null hypothesis (5) the results reveal that the mean amount of information (bits) extracted per stage-I question was significantly (p < .001) higher for the E'CF-16 array than for the NCF-16 array.

**Correlational Analysis**

**Correlation Matrix**

A correlational analysis of the two MFF variables, IQ (CTMM), the two 20-questions informational effectiveness variables and the four 20-questions latency variables was conducted to further test the significance of the relationships that these variables have on each
other (See Table 14). Table 15 contains the means and standard deviations for the variables of Table 14.

**Informational Effectiveness**

The 20-questions latency variables were never significantly correlated with the informational effectiveness variables $E'-E'CF-16$ or $E'-NCF-16$. Also, IQ (CTMM) was not significantly correlated with either of the informational effectiveness variables. In fact, the two informational effectiveness variables ($E'-E'CF-16$ and $E'-NCF-16$) were not significantly correlated with each other. On the other hand, both of the MFF variables (errors and latency) were significantly correlated ($\rho = 0.33$ and $0.31$, respectively) with $E'-E'CF-16$, but neither of the MFF variables were significantly correlated with $E'-NCF-16$. The fact that both accurate MFF performance and high MFF response latency were associated with more effective information gathering on $E'CF-16$, but not $NCF-16$ concurs with and reaffirms the previous findings (Hypotheses 1 and 2).

**20-Questions-Latency**

All four of the 20-questions latency variables posted significant positive intercorrelations ($\rho$, ranged from 0.44 to 0.64) with each other. Moreover, the four 20-questions latency variables were significantly and positively correlated ($\rho$, ranged from 0.27 to 0.39) with MFF latency but none of the 20-questions latency variables were significantly correlated with MFF errors. MFF errors and $E'CF-16$-latency-II, however, posted a mild but nonsignificant correlation ($\rho = -0.24$). Moreover, none of the 20-questions latency variables
TABLE 14
CORRELATION MATRIX

<table>
<thead>
<tr>
<th></th>
<th>E-E'CF-16</th>
<th>E-NCF-16</th>
<th>E'CF-16</th>
<th>NCF-16</th>
<th>IQ(CTMM)</th>
<th>M FF Errors</th>
<th>M FF Latency</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.10</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>E'CF-16</td>
<td>Latency-I</td>
<td>0.03</td>
<td>-0.03</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>E'CF-16</td>
<td>Latency-II</td>
<td>0.15</td>
<td>-0.02</td>
<td>0.60*</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>NCF-16</td>
<td>Latency-I</td>
<td>0.09</td>
<td>0.01</td>
<td>0.51*</td>
<td>0.50*</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>NCF-16</td>
<td>Latency-II</td>
<td>0.11</td>
<td>0.09</td>
<td>0.51*</td>
<td>0.64*</td>
<td>0.44*</td>
</tr>
<tr>
<td>5</td>
<td>IQ(CTMM)</td>
<td>0.15</td>
<td>0.03</td>
<td>0.10</td>
<td>-0.00</td>
<td>0.02</td>
<td>0.06</td>
</tr>
<tr>
<td>6</td>
<td>M FF Errors</td>
<td>-0.33*</td>
<td>-0.07</td>
<td>-0.12</td>
<td>-0.24</td>
<td>-0.09</td>
<td>-0.18</td>
</tr>
<tr>
<td>7</td>
<td>M FF Latency</td>
<td>0.31*</td>
<td>-0.01</td>
<td>0.32*</td>
<td>0.39*</td>
<td>0.35*</td>
<td>0.27*</td>
</tr>
</tbody>
</table>

The correlations with the seventh variable (IQ) were based on 131 subjects whereas all other correlations were based on 139 subjects.

* denotes an overall significance level of p < .05; in short, all * correlations were pairwise significant at or beyond p < .001. Due to the skewness and nonparametric nature of some of the distributions, the Spearman Rank correlations were used.
TABLE 15

MEANS AND STANDARD DEVIATIONS OF THE VARIABLES OF TABLE 14

<table>
<thead>
<tr>
<th></th>
<th>E-CF-16</th>
<th>E-CF-16</th>
<th>NCF-16</th>
<th>NCF-16</th>
<th>IQ (CTMM)</th>
<th>MFF</th>
<th>MFF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Latency-I</td>
<td>0.93</td>
<td>0.09</td>
<td>0.87</td>
<td>0.08</td>
<td>3.4</td>
<td>2.6</td>
<td>4.5</td>
</tr>
<tr>
<td>Latency-II</td>
<td>3.7</td>
<td>2.8</td>
<td>4.5</td>
<td>3.3</td>
<td>119.6</td>
<td>9.9</td>
<td>16.8</td>
</tr>
<tr>
<td>Errors</td>
<td>16.8</td>
<td>8.4</td>
<td>37.8</td>
<td>30.8</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The latencies were measured in seconds while the E scores were measured in bits.
were significantly correlated with IQ (CTMM).

**MFF Variables and IQ**

Neither of the MFF variables (errors and latency) were significantly correlated with IQ (CTMM). MFF latency and IQ (CTMM), however, exhibited a mild but nonsignificant correlation ($\rho = 0.23$). The two MFF variables were correlated at -0.70.

**Summary and Discussion**

There were two significant findings. First, as predicted, there was more effective question asking (i.e., the hierarchical organization of a participant's line of questioning was more in accord with the optimal "half by half. . ." principle) on E'CF-16 than on NCF-16. Second, as predicted on the E'CF-16 array, the extreme reflective subjects were more effective question askers than their fast-accurate, slow-inaccurate and extreme impulsive counterparts. The "optimal strategy rate" data further attests to the fact that the extreme reflectives were the superior and dominant tempo group on the E'CF-16 array. Moreover, the inclusion of the non-extremes (impulsives-reflectives) in the ordinary tempo groups simply diluted (.05 < p < .1) the otherwise significant differences.

Overall, the correlational analysis confirmed the experimental findings. The correlational analysis, moreover, attests to the stability of response latency, both within the 20-questions task and across the two tasks (MFF and 20-questions game).
The above findings could be confounded by the amount of specific hypothesis guessing (HS). Perhaps the extreme reflectives were more effective question askers simply because the members of the other tempo groups used substantially more specific hypothesis seeking (HS) questions. If the differences in effective question asking can be attributed to a differential deployment of HS questioning, then the contention that—extreme reflective subjects hierarchically organized their line of questioning more in accord with the optimal "half by half. . ." principle—is of a rather technical and superficial nature. Likewise, differences in specific hypothesis seeking (HS) versus constraint seeking, rather than the facilitating effect of the 2^6 proportioning, could account for the more effective questioning on E'CF-16. Table 16 gives the percent^7 of Stage-I specific hypothesis seeking questions for the two arrays and various tempo groups.

TABLE 16
PERCENT OF HS QUESTIONS OCCURRING IN STAGE-I ARRAYS BY TEMPO GROUPS

<table>
<thead>
<tr>
<th></th>
<th>Impulsives</th>
<th>Impulsives</th>
<th>Slow-</th>
<th>Fast-</th>
<th>Reflectives</th>
<th>Reflectives</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Extreme</td>
<td>Nonextreme</td>
<td>Inaccurate</td>
<td>Accurate</td>
<td>Nonextreme</td>
<td>Extreme</td>
</tr>
<tr>
<td>E'CF-16</td>
<td>10.0%</td>
<td>9.5%</td>
<td>2.6%</td>
<td>2.6%</td>
<td>6.8%</td>
<td>4.3%</td>
</tr>
<tr>
<td>NCF-16</td>
<td>9.3%</td>
<td>8.7%</td>
<td>1.9%</td>
<td>2.1%</td>
<td>3.8%</td>
<td>3.7%</td>
</tr>
</tbody>
</table>

^7 On the 16-item arrays used in this investigation, the number of Stage-I questions did not markedly differ across arrays or tempo groups. Hence, the percents of Table 16 are indicative of the number of HS occurrences.
The data of Table 16 indicates that across all tempo groups there was slightly less specific hypothesis guessing (HS) on NCF-16 than on E'CF-16. Thus, the more effective question asking on E'CF-16 as opposed to NCF-16 can not be attributed to different rates of specific hypothesis seeking. In fact, if anything, the specific hypothesis seeking rates of Table 16 would suggest more effective question asking on NCF-16 than on E'CF-16.

The data of Table 16 further reveals that, regardless of array, impulsive subjects evidenced slightly higher rates of specific hypothesis seeking than reflective subjects, and reflective subjects, in turn, exhibited slightly higher rates of HS questioning than their fast-accurate and slow-inaccurate peers. However, the data of Table 16 indicates that, overall, the rate of HS questioning was rather insignificant and hence it is doubtful if these slight fluctuations in specific hypothesis seeking markedly altered E (i.e., informational effectiveness) scores. Moreover, this position can be further supported by the following argument: if the rate of specific hypothesis seeking markedly influenced effective question asking, then the fast-accurate and slow-inaccurate subjects should have been superior to impulsive subjects (both extremes and nonextremes). This, however, is contrary to the findings reported in Table 3.

Only two of the 139 respondents initiated their information gathering with an HS question. Thus, since most of the HS questioning occurred toward the end of Stage-I, the effect of the HS questioning was further minimized.
In sum, the view that the significant findings (array and tempo group) of this investigation could be confounded by the amount of specific hypothesis guessing can be dismissed. Moreover, the non-significant correlation ($\rho = 0.15$) between IQ and $E-E'CF-16$ is sufficient to discount IQ as a confounding factor.

Lastly, one issue of interest remains. The correlational analysis revealed that:

(1) both accurate MFF performance and lengthy MFF response times were associated with effective question asking on E'CF-16;

(2) a preference for reflective delay on the MFF task was associated with lengthy E'CF-16-latency-II;

(3) but, a preference for longer E'CF-16-latency-II was not significantly related to more effective question asking on E'CF-16.

In view of the above, the information of Table 17 is of interest. In short, on the MFF task the extreme reflectives were the highly latent subjects while the extreme impulsives were the excessively fast responders. However, on E'CF-16-latency-II, the overall trend remained the same but the extreme reflectives shifted from a highly latent to a midrange latency group.

These results help to explain the rather low ($\rho = 0.15$) correlation between $E-E'CF-16$ and E'CF-16-latency-II. In other words, with a number of the highly effective E'CF-16 question askers (namely, the extreme reflectives) posting only midrange response latencies on E'CF-16-latency-II, the degree of positive relationship between $E-E'CF-16$ and E'CF-16-latency-II was severely attenuated.
<table>
<thead>
<tr>
<th></th>
<th>Percentage</th>
<th>Percentage</th>
<th>Percentage</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>% E. Reflectives</td>
<td>% E. Impulsives</td>
<td>% E. Reflectives</td>
<td>% E. Impulsives</td>
</tr>
<tr>
<td><strong>MFF Latency</strong></td>
<td>0.0%</td>
<td>100%</td>
<td>10.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>90.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td><strong>E'CF-16 Latency-II</strong></td>
<td>0.0%</td>
<td>68.4%</td>
<td>80.0%</td>
<td>15.8%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>20.0%</td>
<td>15.8%</td>
</tr>
</tbody>
</table>

**TABLE 17**

COMPARISONS OF HOW THE EXTREMES (REFLECTIVE-IMPULSIVE) DISTRIBUTED ABOUT THE MEAN ON TWO MEASURES OF LATENCY
It appears that accurate performance on the MFF task required a rather time consuming search and elimination procedure while effective information gathering on E'CF-16 requires only a few moments consideration (See Table 15). Moreover, it further appears that the extreme reflectives are not just highly latent per se, but have the flexibility to adjust their response latency to the task requirements.
CHAPTER V

SUMMARY, CONCLUSIONS AND IMPLICATIONS

Summary

This study investigated the differential effect of conceptual tempo on the effectiveness of the information gathering strategies of sixth grade boys engaged in two variants (E'CF-16 and NCF-16) of the 20-questions game.

Conceptual tempo was operationalized by applying a split median analysis to the total errors and mean latency of initial response on the Adult Adolescent version of the Matching Familiar Figures Test (Kagan et al., 1964). The error-latency split median analysis classified participants as: fast-inaccurate (impulsives), fast-accurate, slow-accurate (reflectives) and slow-inaccurate.

The data were collected by a two-phase process. In Phase I, the MFF test was individually administered to 140 sixth-grade boys from a suburban school system. The Phase I MFF response style categorization produced 52 impulsives, 50 reflectives, 19 fast-accurate and 19 slow-inaccurate.

In Phase II, the 20-questions procedure was individually administered to each participant. Each participant played two games on each array--NCF-16 and E'CF-16.
In order to have equal cell sizes, 18 subjects were randomly selected from each of the four conceptual tempo categories. Since two arrays were used, there were two possible orders of presentation—NCF-16 and then E'CF-16, or vice versa. Participants were randomly assigned to an order of presentation so that exactly half of the 18 respondents of each conceptual tempo category received one order of presentation while the remaining members of that response style received the other order of presentation.

E'CF-16 was a 4x4 color-form array consisting of four equally distributed bivariate dimensions. Thus, E'CF-16 had the convenient $2^4$ proportioning. Moreover, it was theorized that the $2^4$ proportioning would guide a subject's performance toward the optimal "half by half by . . ." strategy and thereby facilitate effective information gathering. In short, the E'CF-16 array was characterized as a rather easy array with a few "pitfalls." The "pit-falls" were: 1) redundancy error, 2) conjunctive error, and 3) row-column error.

NCF-16 was a 4x4 color-form array consisting of two quadri-variate dimensions and one trivariate dimension. For each of the three dimensions, the variate levels were nonequally distributed. Hence each dimension contained some questions that were informationally "better" than others. NCF-16 was also constructed away from the $2^4$ proportioning by placing the highest potential for optimal solution on the attribute (letter) of least dimensional saliency.

Although NCF-16 was designed to make optimal solution difficult, one could attain a relatively effective Stage-I information gathering strategy by initially questioning on the most numerous level of a.
dimension and then progressing toward the less numerous levels. However, this strategy yielded an optimal solution only in the letter dimension. Clearly, this "big to little" strategy is of a rather spurious nature. In all, an effective performance on the NCF-16 array (quite likely) required more concerted effort on the part of the respondent.

The experiment involved a four tempo by two orders of presentation design. However, there was a concern over the distinctiveness among the four conceptual tempo groups. Moreover, it was deemed desirable to enhance the distinctiveness among tempo groups by use of an extreme groups design. The extreme reflectives were those participants who were below the 25th percentile on MFF errors and above the 75th percentile on mean latency of initial response whereas the extreme impulsives were those respondents who were below the 25th percentile on mean latency of initial response and above the 75th percentile on total errors. This schema produced no extreme fast-accurates or slow-inaccurates. The fast-accurates and slow-inaccurates were compared with both the extreme reflective-impulsive groups and the ordinary reflective-impulsive groups.

In all there were nine experimental variables: two informational effectiveness variables, four 20-questions latency variables, IQ (CTMM) and two MFF variables.

The data were then analyzed: 1) with 4x2 univariate analysis of variance for extreme groups and for ordinary groups, on each of the two informational effectiveness variables; 2) with a nonparametric 4x2 randomized block test; and 3) by correlational analysis of the nine experimental variables.
Also, $E^{'}ECF-16$ and $E-NCF-16$ were appropriately collapsed into a single informational effectiveness variable $\bar{E}$, and used as the criterion variable in the ordinary groups by grid univariate analysis of variance procedure.

Conclusions

The problem studied in this investigation was: the differential effect of conceptual tempo on the effectiveness of information gathering strategies. This overall problem was divided into the following predictions and/or subproblems:

(1) It was predicted that reflective subjects would be more likely, than their conceptual tempo counterparts, to consider the consequences of their 20-questions (both arrays) queries and thereby employ a line of questioning in which the hierarchical organization more closely approximated the optimal "half by half. .." principle. In short, it was predicted that the reflective subjects $\bar{E}$ scores (for both arrays--$E^{'}CF-16$ and NCF-16) would be significantly higher than those of the other tempo groups.

(2) It was predicted that the $2^4$ proportioning of array attributes would facilitate the use of the optimal "half by half. .." principle. In other words, it was predicted that the $2^4$ proportioned color-form array ($E^{'}CF-16$) will be "easier" (i.e., the $E^{'}CF-16$ array will post significantly higher $\bar{E}$ scores) than the non-$2^4$-proportioned color-form array (NCF-16).
(3) The presence of an "order of presentation" effect was tested for both arrays.

(4) The presence of a conceptual tempo-order of presentation interaction was tested for both arrays.

(5) Lastly, it was predicted that any differential effect due to conceptual tempo would be more prevalent at the extremes of the reflective-impulsive dimension.

In all, subproblems two through four dealt with factors that were confounding or orthogonal to subproblem one while subproblem five dealt with the locus of strength of the resulting tempo effects. Thus, in order to summarize and discuss the main findings of the study, the subproblems were organized under the following headings:

(1) Confounding Factors

(2) Differential Effect of Conceptual Tempo on Effectiveness of Information Gathering

The conclusions and interpretations cannot, except in strictly speculative sense, be carried beyond the following delimitations:

(1) Above average to high IQ (Mean = 119.6 and Standard Deviation = 9.9) sixth grade boys from a suburban school system served as the sample.

(2) Tasks consisted of two variants of the 20-questions game.
Confounding Factors

Hypothesis (5) examined the question: Does grid (E'CF-16 vs. NCF-16) have a differential effect on informational effectiveness?

Analysis of variance tests have shown (see pp. 79-81, Table 13 and Graph 5) that the mean amount of information (bits) extracted per Stage I question was significantly higher for the E'CF-16 array than for the NCF-16 array. The rate of specific hypothesis seeking (HS) data revealed that specific hypothesis guessing did not account for the less effective question asking on NCF-16 (see Table 16 and discussion on p. 88). This is as predicted, and hence supports the theorized position that, although both grids contained highly salient (i.e., readily available) attributes for categorizing and partitioning the stimulus array, the $2^4$ proportioning of E'CF-16 would facilitate more effective information gathering strategies than the non-$2^4$-proportioned NCF-16 array.

Moreover, the optimal solution rate data revealed that on E'CF-16 35.3% of the 139 phase-II participants evidenced an optimal solution strategy while only 0.7% of the phase-II respondents exhibited an optimal performance on NCF-16. This is further support for the previously theorized position. Thus, from an informational effectiveness point of view, one can conclude that NCF-16 is the more difficult of the two arrays.

Hypotheses (3) and (4) examined the question: Does order of presentation have a differential effect on informational effectiveness?

Analysis of variance tests (for both arrays) revealed no significant order of presentation effects nor any significant conceptual tempo-
order of presentation interaction effects (see pp. 70-77 and Tables 5, 7, 9 and 11).

The correlational analysis (Table 14, p. 85) revealed that IQ (CTMM) was not significantly associated with effective question asking on either of the two 20-questions arrays (E'CF-16 or NCF-16). Moreover, effective question asking on E'CF-16 was not significantly correlated with effective question asking on NCF-16. Thus, it appears that not only is NCF-16 the more difficult of the two arrays but, effective information gathering on the two arrays (E'CF-16 and NCF-16) is, at least performatively speaking, fundamentally different in nature. In summary and conclusion, it appears that of all the extraneous factors considered, only the issue of array proportionality remains a serious threat to confound the conceptual tempo prediction.

**Differential Effect of Conceptual Tempo on Effectiveness of Information Gathering**

Hypotheses (1) and (2) examined the question: Does conceptual tempo have a differential effect on the effectiveness of one's information gathering strategy?

The Hettmansperger test (Tables 2 and 3, discussion pp. 65-66), revealed that on the E'CF-16 array, extreme reflectives were significantly more effective information gatherers than their slow-inaccurate, fast-accurate and extreme impulsive counterparts while the analysis of variance and Hettmansperger tests revealed no significant tempo effects (for either extreme or ordinary groups) on the NCF-16 array. On E'CF-16, the overall trend was in the direction of reflective subjects being more effective information gatherers than their
conceptual tempo counterparts. However, with the ordinary groups these differences were not significant at $p < .05$. The data on optimal strategies (see p. 69) concurred with the above findings. Also, the data on rate of specific hypothesis seeking (Table 16 and discussion on p. 88) dispelled the notion that the tempo effect reported on E'CF-16 could be attributed to differential rates of specific hypothesis seeking.

How does one interpret the finding of a significant tempo effect on E'CF-16 but no accompanying tempo effect on NCF-16? The Confounding Factors section (pp. 98-99) showed that of all the confounding variables considered, only the proportionality of array attributes had a significant effect on informational effectiveness. In short, the $2^4$ proportioning of E'CF-16 educed more effective question asking on E'CF-16 than the non-$2^4$-proportioning did on NCF-16. Thus, NCF-16 was regarded as the more difficult of the two arrays. Perhaps, the difficulty factor confounds the tempo effect? This issue requires further examination.

The 35.3% optimal performance rate on E'CF-16 and the 0.7% optimal performance rate on NCF-16 suggests that the participants did not fully grasp the "half by half. . ." principle. Surely, persons who fully grasped and had command of such a principle would be in position to act on the arrays rather than being influenced by them. Also, the data from the interviews conducted at the completion of the 20-questions tasks suggested that the children did not comprehend the more generic aspects of the "half by half. . ." principle.
These conclusions are in agreement with Ault's (1973) remarks:

Informal observations suggest that few children, even at the fifth grade, knew that the optimal strategy was to divide the array in half with each question. Rather, they appeared to believe that a particular type of question was appropriate to the game, but specific questions fitting that type were asked without regard to efficiency. The most common example was the perceptual question about color. The Ss did not ask about the more frequently represented colors before the less frequent ones, nor did they appear to notice that some color questions had only one referent and hence were pseudo questions. It seemed, instead, that Ss had a set to ask about colors and persisted until a yes response was obtained.¹

Thus, the more effective question asking on E'CF-16 and, in particular, the extreme reflectives highly effective performance on E'CF-16 cannot be unequivocally attributed to a better understanding of the "half by half. . ." principle. Instead, it appears that the 2⁴ proportioning of E'CF-16 rather incidentally sets up the "half by half. . ." questioning.²

This is not to imply that the E'CF-16 respondents were not "thinking"; but is merely to assert that it is doubtful if they were employing (in the full sense of the word) the "half by half. . ." principle. Surely, one can be very shrewd in dealing with the

¹The participants of the current investigation were highly intelligent (mean IQ=119.6) sixth graders and may have been slightly more aware of the "half by half. . ." principle but, in all, Ault's remarks generally concur with the informal observations of the current study. Also, it is important to note that Ault's arrays lacked the convenient 2⁴ proportioning.

²Perhaps this explains why effective performance on E'CF-16 did not carry over to NCF-16. Note: This type of carryover would have induced an order effect on NCF-16 and no such order effect was exhibited.
specifics of a situation and yet fail to comprehend the more generic consequences. The following excerpts from the interview sessions help to illustrate this point:

If (on E'CF-16) you have found the shape and color, which question should you ask next: (1) Is it capital H? or (2) Is it the letter h? Or (3), doesn't it make any difference?

To this question children would frequently respond:

I'd ask the letter h, because if you ask capital H and the answer is "no" then you won't know which of the three it is. If you ask letter h--well, that divides things in half and I'll know whether it's these two things with letter h or these two things with letter b.

Yet this same person when questioned about aspects of NCF-16 would not invoke the "half by half. . ." principle.

Also, (on E'CF-16) several of the extreme reflectives voluntarily remarked:

I thought of asking about rows but didn't because it could take three questions to find the row and then there would still be four things left. . .if I ask color, shape and letter, three questions will get it down to two things.

Even though the participants failed to comprehend the "half by half. . ." principle, it appears that on E'CF-16 a number of respondents considered the consequences of their queries and were thereby able to distinguish more effective lines of questioning from less effective ones. Furthermore, the results suggest (as predicted) that the extreme reflectives were more apt (than the other tempo groups) to consider the consequences of their actions and were thereby more successful (than the other tempo groups) in avoiding the "pitfalls" of E'CF-16.
NCF-16, on the other hand, had the highly salient attributes but lacked the $2^4$ proportioning. Hence, on NCF-16 an optimally effective strategy was not readily available and therefore distinctions between effective and ineffective lines of questioning were dependent on a command of the "half by half..." principle. Perhaps, on NCF-16 the general failure to comprehend the "half by half..." principle tended to impair any benefits to be derived from a preference for reflective delay.

In conclusion, the presence of a significant tempo effect on E'CF-16 with no accompanying tempo effect on NCF-16 can be explained (as in the previous two paragraphs) by the absence of the "half by half..." principle in conjunction with the facilitating effect of the $2^4$ proportioning. This conclusion, however, requires some minor qualifications. These qualifications stem from the fact that E'CF-16 cannot represent all $2^4$ proportioned arrays but is, indeed, a particular array with the $2^4$ proportioning. Likewise, NCF-16 is a particular non-$2^4$-proportioned array.

E'CF-16 has been characterized as a rather simple (i.e., highly salient attributes and $2^4$ proportioning) 20-questions array with a few "pitfalls" (e.g., redundancy error, row-column error and conjunctive error). The redundancy and row-column errors would be "pitfalls" on any $2^4$ proportioned array. The conjunctive error, however, is of special interest. Only one out of 139 subjects committed a conjunctive error of the form: "Is it a red square?" All other conjunctive errors involved the letter-case attributes; e.g.,
"Is it a capital H?" This suggests that the respondents did not seek out a conjunctive approach but, instead, were "tricked" into it by the "natural" tendency to collapse the letter-case dimensions. Thus, it appears that on a $2^4$ proportioned array one could maximize (minimize) the informational effectiveness benefits to be derived from a preference for reflective delay by appropriately altering the number and nature of the "pitfalls" that are built into the particular array. In short, the $2^4$ proportioning provides the task do-ability while the "pitfalls" bring out the tempo effect.

The majority of the NCF-16 information gathering choices ranged between questions which classified one fourth to three fourths of the array objects. Perhaps, the "pitfalls" and lacunae of NCF-16 were too subtle to distinguish differences between tempo groups. If the non-$2^4$-proportioned array were set up so that the distinctions between "good" moves and "poor" moves were more exaggerated then the task may be more do-able (i.e., less dependent on a firm command of the "half by half..." principle) and provide the type of pitfalls and lacunae that would yield a significant tempo effect.

The issue of task do-ability is an important one. Essentially, it is the issue raised by Kagan (1965) when he declares that the benefits of a preference for reflective delay are maximal when the respondents have acquired (but not overlearned) the "rudimentary skills" necessary for the performance of the task.

Bush and Dweck (1975) proposed a model to predict the relative speed and accuracy of reflectives and impulsives as a function of the required task strategy. Via their model Bush and Dweck theorized
that, "...on an easy task, where a mature strategy is not necessary, that is, when the strategy required for optimal performance is simple and straightforward, one would not expect any performance differences between reflectives and impulsives in either rate or accuracy."

However, in the current investigation the tempo effect favoring the extreme reflectives occurred only on E'CF-16--the easier of the two arrays. Perhaps, E'CF-16 was not easy enough to meet Bush and Dweck's notion of "easy"; at any rate, the NCF-16 results suggest that task difficulty can also prevent a tempo effect. Here Kagan's (1965) "rudimentary skills" requirement is beneficial in that it brings the participants (relevant) skills and capacities into the determination of task difficulty.

The finding that extreme reflectives were extremely slow responders on the MFF task but displayed only mid-range latencies on E'CF-16-latency-II (extreme impulsives were among the extremely fast responding on both tasks, p. 91) suggests that the reflectives are not slow responding per se but, indeed, modify their response times to meet the situational demands. This concurs with Bush and Dweck's (1975) conclusion that reflectives have the flexibility to take the requirements of the current task into account and allow themselves to perform accordingly. These findings suggest that, perhaps, a preference for reflective delay has more to do with strategy preference than latency per se.

3Although the 24 proportioned E'CF-16 was basically an easy array, it did contain some rather deceptive pitfalls.
Lastly, the findings reported in Tables 2 and 3 in conjunction with the optimal strategy rate data (pp. 65-66) demonstrated that in regard to effective information gathering on E'CF-16 the extreme reflectives were the superior and dominant tempo group. Also, the E'CF-16 ordinary tempo group results (Table 4) revealed that the inclusion of nonextremes (reflectives-impulsives) attenuated the tempo effect. In all, these findings confirm the prediction that the existent tempo effect was more prevalent at the extremes of the reflective-impulsive dimension.

Summary

In regard to effective information gathering the main findings were:

(1) No significant order of presentation effect for either array.

(2) No significant order of presentation-conceptual tempo interaction for either array.

(3) No significant relationship (on either array) with IQ (CTMM).

(4) As predicted: a significant proportionality effect with NCF-16 (the non-$2^4$-proportioned array) being more difficult than E'CF-16 (the $2^4$ proportioned array).

(5) As predicted: a significant tempo effect on E'CF-16 favored the extreme reflectives over their conceptual tempo counterparts. However, the predicted tempo effect favoring reflectives) failed to materialize on NCF-16.
Further, investigation revealed that (potential) differential rates of specific hypothesis seeking did not confound the significant findings (4 and 5 above). Also, informal interviews indicated that (overall) the participants did not fully comprehend the advantages of the "half by half. . ." principle.

The result that a significant tempo effect occurred on only one (E'CF-16) of the two arrays was explained by the absence of the "half by half. . ." principle and the proportionality (i.e., difficulty) factor. In short, it was theorized that: on NCF-16, the proportionality of attributes in conjunction with the absence of the "half by half. . ." principle blurred the distinctions between effective and ineffective information gathering choices whereas the 2^4 proportioning of E'CF-16 establishes an optimal information gathering situation in which the participants had the capability (provided they considered the alternatives) to distinguish effective questioning from ineffective questioning.

The preference for reflective delay on the MFF task showed stability and persisted across the 20-questions latency measures. However, the results also revealed that the extreme reflectives can modify their response latencies to meet the requirements of the current 20-questions task. This suggests that, perhaps, the preference for reflective delay is more a strategy preference rather than a latency preference per se.

Lastly, and as predicted, the results showed that the E'CF-16 tempo effect was more prevalent at the extremes of the reflective-impulsive dimension. In fact, in regard to effective or optimal
information gathering on E'CF-16, the extreme reflectives were the dominant tempo group.

**Implications**

The implications were organized under three headings:

1. Implications for Research Methodology
2. Implications for Further Research
3. Implications for Academic Performance

**Implications for Research Methodology**

The result that the existent (E'CF-16) tempo effect was more prevalent at the extremes of the reflective-impulsive dimension, indicates the desirability of an extreme groups design for conceptual tempo research. In fact, strictly speaking, the E'CF-16 tempo effect only existed within the extreme groups design. Thus, failure to use extreme groups would have circumvented the findings of significant differences.

Further, and most importantly, the significant differences did not merely exist between the two extreme groups (i.e., extreme reflectives versus extreme impulsives); instead, the findings revealed that in regard to effective or optimal information gathering on E'CF-16, the extreme reflectives were the dominant tempo group. Hence, the loss from a failure to use extreme groups can be substantial.

A number of researchers (e.g., Ault, 1973; Callahan and Senh, Lata, Pass, 1971) have reported, "nonsignificant differences in the predicted direction," and one must wonder what an extreme groups
design could have contributed. In all, an extreme groups design can be highly recommended for conceptual tempo research.

Implications for Further Research

The implications for further research will be restricted to the conceptual tempo–informational effectiveness–array proportionality (i.e., difficulty) issue. Under this restriction the immediate research questions are:

1. If a second $2^4$ proportioned array is formed by removing the letter-case conjunctive error, will the newly formed array (a) be less difficult than E'CF-16 and (b) show no significant tempo effect?

2. If a second non-$2^4$-proportioned color-form array is constructed to exaggerate (more than NCF-16) the informational ineffectiveness of initial color-shape queries, will this newly constructed array evidence a tempo effect favoring the reflectives?

The results of the current investigation suggest that persons with a strong preference for reflective delay are more apt to spot strategic deficiencies or pitfalls (provided the subtlety of the deficiencies or lacunae does not exceed the respondent's performative capabilities) in a proposed course of action. Research question (1) purports to test the "seeing deficiencies" notion by showing that the removal of deficiencies or pitfalls attenuates the tempo effect whereas research question (2) purports to test the "seeing deficiencies" notion by demonstrating that a tempo effect favorable to reflective subjects can occur on a non-$2^4$-proportioned array.
If the "seeing deficiencies" notion can be further substantiated it could provide a link between conceptual tempo and a dimension that Guilford (1965 and 1971) and his collaborators have identified as "sensitivity to problems." It appears that epistemic behavior and general problem solving would be highly dependent on responsiveness to problematic discrepancies and lacunae. In this light Berlyne (1961, Preface of Conflict, Arousal and Curiosity) points out that, "The process of raising problems is often hard to separate from the process of suggesting solutions to problems."

**Implications for Academic Performance**

The implications for academic performance stem from the "seeing deficiencies" notion. Thus, for discussion purposes, it will be assumed that the "seeing deficiencies" notion has survived the scrutiny of further testing. Also, it should be noted that, by its very nature, this section takes on a more speculative tone. Hence, the issues discussed herein require further research--both in a controlled experimental atmosphere and in a field setting.

The "seeing deficiencies" notion will be cast in the following problematic setting:

A first pipe fills a given tank in six hours while a second pipe fills the same tank in three hours. If the pipes run simultaneously, how long will it take both pipes together to fill the tank?

Since the issue is "seeing deficiencies" in proposed solutions, the focus of attention will be directed toward typical "errors"
rather than "how to get the answer." Four and a half hours is an often reported incorrect answer, i.e., the person finds the average. Note, however, the first pipe fills the tank in three hours and hence plausible solutions must lie between zero and three hours. In short, the problem requirements (if considered) immediately and rather obviously suggest a range of plausible solutions. In all, it seems quite reasonable to conjecture that a strong preference for reflective delay would be beneficial to the avoidance of such (e.g., the "averaging") pitfalls.

The "sensitivity to problems" dimension comes under the general heading of critical thinking. Critical thinking, moreover, is often stated as a general aim of education. Thus, it seems worthwhile to test the conceptual tempo-seeing deficiencies notion in a variety of social, political, economic and governmental settings as well as testing this notion in mathematical and scientific settings.
APPENDIX

MFF SAMPLE ITEM
BIBLIOGRAPHY


