EFFECTS OF STIMULUS DIMENSIONALITY
AND OTHER STIMULUS CHARACTERISTICS
ON CLASSIFICATION PERFORMANCE
IN CHILDREN OF AGES 3 TO 7 YEARS

by

Donna Marilyn Scribner
B.H.E., University of British Columbia, 1961

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF ARTS

in the Department
of
Educational Psychology

We accept this thesis as conforming to the
required standard

THE UNIVERSITY OF BRITISH COLUMBIA
August, 1974
In presenting this thesis in partial fulfilment of the requirements for
an advanced degree at the University of British Columbia, I agree that
the Library shall make it freely available for reference and study.
I further agree that permission for extensive copying of this thesis
for scholarly purposes may be granted by the Head of my Department or
by his representatives. It is understood that copying or publication
of this thesis for financial gain shall not be allowed without my
written permission.

Department of Educational Psychology

The University of British Columbia
Vancouver 8, Canada

Date August, 1974
Abstract

The problem studied in this investigation was the effects of stimulus dimensionality, other stimulus characteristics and the age of subjects on free classification performance. The theoretical basis was earlier research (Inhelder & Piaget, 1964) and the experimental procedures were adapted from a test developed to assess conceptual functioning in brain damaged children (Graham-Ernhart Block Sort Test, 1963) GEBST.

Ninety children, 18 from each of the following age groups; preschool three year olds, preschool four year olds, kindergarten, grade one and grade two were subjects in this investigation. There were three sets of stimulus objects; three dimensional styrofoam objects, two dimensional flat styrofoam objects and photographs of the three dimensional objects. Each stimulus set required subjects to match or group stimuli which differed in respect to three characteristics; color (yellow, red and blue), form (circle, square and triangle), and size (large, medium and small).

The GEBST procedures were followed with minor changes in the order of task presentation and in the scoring. Two difficulty levels in the GEBST were used and each level contained three trials. Level III stimuli differed on one characteristic and were constant in respect to the remaining characteristics. Level IV stimuli differed on two characteristics simultaneously while the third characteristic was held constant. Dimensionality of the test material was randomly assigned with six subjects in each age group receiving the GEBST in each dimension. A subject received only one dimension but all three characteristics.
Two separate analysis of variance procedures were performed, one for each level on the GEBST. Dimensionality, characteristics, age as a categorical variable and their interactions were assessed with analysis of variance. Age as a continuous variable, sex and their interactions were assessed with two least square regression analyses, one for each difficulty level of the GEBST. Unequal numbers of each sex for some age groups necessitated analysis of the sex variable by least squares regression analysis.

Age was the most significant factor affecting performance on a free classification task. The age range of this investigation seemed adequate to test the performance of preconceptual classification behavior as described by Inhelder & Piaget (1964). The sex of subjects did not significantly affect classification performance, a finding supported by most research in this area. The characteristic of the stimulus object was a significant factor affecting classification performance. Size appeared to be a much more difficult characteristic to base classification on than color or form particularly with the youngest subjects.

Stimulus dimensionality was not found to be a significant factor affecting classification performance. This was contrary to research in this area and also the hypotheses made in this investigation. Two explanations seemed plausible for the results occurring in this study: the scoring system of the GEBST did not sufficiently reflect the performance of the tested subjects, and the choice paradigm of Level IV, implicitly in the design of the GEBST, obscured some of the variability of stimulus dimensionality.

Further research in the area of stimulus dimensionality in a
free classification task seems necessary as the results of this investigation did not support other investigations in this area. It is also an area which has not been fully studied and stimulus dimensionality has great practical applicability to the study of young children's performance on conceptual tasks.
# Table of Contents

## Chapter 1  Introduction

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Statement of the Problem.</td>
<td>1</td>
</tr>
<tr>
<td><strong>Review of Literature</strong></td>
<td></td>
</tr>
<tr>
<td>Theoretical Aspects of the Problem</td>
<td>1</td>
</tr>
<tr>
<td>Validation Research on Classification Theories</td>
<td>4</td>
</tr>
<tr>
<td>Conceptual Functioning Assessment and Classification Tasks</td>
<td>7</td>
</tr>
<tr>
<td>Training Procedures and Classification Research</td>
<td>11</td>
</tr>
<tr>
<td>Subject Characteristics and Classification Research</td>
<td>13</td>
</tr>
<tr>
<td>Stimulus Attributes and Classification Research</td>
<td>15</td>
</tr>
<tr>
<td>Procedures in Classification Research</td>
<td>23</td>
</tr>
</tbody>
</table>

## Chapter 2  The Research Problem

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Purpose</td>
<td>27</td>
</tr>
<tr>
<td>Operational Definitions</td>
<td>27</td>
</tr>
<tr>
<td>Hypotheses</td>
<td>28</td>
</tr>
</tbody>
</table>

## Chapter 3  Methods

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subjects</td>
<td>30</td>
</tr>
<tr>
<td>Stimuli</td>
<td>32</td>
</tr>
<tr>
<td>Procedure</td>
<td>33</td>
</tr>
<tr>
<td>Scoring</td>
<td>36</td>
</tr>
<tr>
<td>Design</td>
<td>37</td>
</tr>
<tr>
<td>Statistical Analysis</td>
<td>38</td>
</tr>
</tbody>
</table>

## Chapter 4  Results

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
</table>

## Chapter 5  Summary and Conclusions

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
</table>

## References

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
</table>

## Appendix

<table>
<thead>
<tr>
<th>Section</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>A Graham-Ernhart Block Sort Test Manual</td>
<td></td>
</tr>
<tr>
<td>B Consent Form Letter</td>
<td></td>
</tr>
<tr>
<td>C Pictorial Representational Stimuli</td>
<td></td>
</tr>
</tbody>
</table>
# List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Distribution of Subjects by Age and Sex</td>
<td>31</td>
</tr>
<tr>
<td>2 Stimulus Objects</td>
<td>34</td>
</tr>
<tr>
<td>3 Summary of Analysis of Variance of Classification Performance on Level III</td>
<td>43</td>
</tr>
<tr>
<td>4 Summary of Analysis of Variance of Classification Performance on Level IV</td>
<td>47</td>
</tr>
<tr>
<td>5 Mean Comparison of Stimulus Characteristics on Level IV using the Bonferroni Test with a 95% Confidence Interval</td>
<td>50</td>
</tr>
<tr>
<td>6 Mean Comparison of Stimulus Characteristics for the 3 Year Age Group on Level IV using the Bonferroni Test with a 95% Confidence Interval</td>
<td>54</td>
</tr>
<tr>
<td>7 Summary of Stepwise Regression of Independent Variable Predicting Classification Performance on Level III</td>
<td>59</td>
</tr>
<tr>
<td>8 Summary of Stepwise Regression of Independent Variable Predicting Classification Performance on Level IV</td>
<td>60</td>
</tr>
</tbody>
</table>
# List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Diagram of the Research Design</td>
<td>39</td>
</tr>
<tr>
<td>2</td>
<td>Mean Classification Performance Scores with the Effects of Age on Level III</td>
<td>44</td>
</tr>
<tr>
<td>3</td>
<td>Mean Classification Performance Scores with the Interaction of Stimulus Dimensionality and Stimulus Characteristics on Level III</td>
<td>45</td>
</tr>
<tr>
<td>4</td>
<td>Mean Classification Performance Scores with the Effects of Age on Level IV</td>
<td>48</td>
</tr>
<tr>
<td>5</td>
<td>Mean Classification Performance Scores with the Effects of Characteristics on Level IV</td>
<td>49</td>
</tr>
<tr>
<td>6</td>
<td>Preferred or First Selected Characteristics for the Five Ages Tested on Level IV</td>
<td>51</td>
</tr>
<tr>
<td>7</td>
<td>Mean Classification Performance Scores with the Interaction of Age and Stimulus Characteristics on Level IV</td>
<td>53</td>
</tr>
<tr>
<td>8</td>
<td>Mean Classification Performance Scores with the Effects of Dimensionality on Level IV</td>
<td>56</td>
</tr>
<tr>
<td>9</td>
<td>Mean Classification Performance Scores with the Interaction of Stimulus Dimensionality and Stimulus Characteristics on Level IV</td>
<td>57</td>
</tr>
</tbody>
</table>
Chapter 1
Introduction

Statement of the Problem

Research with classification tasks is an old problem that has been done many times. Although a large number of studies have been conducted, few have investigated the effect of the many subject and stimulus characteristics on classification performance. The major concern of this investigation was to assess effects of stimulus dimensionality, stimulus characteristics, and subject's age on free classification performance. The theoretical basis was earlier classification research (Inhelder & Piaget, 1964) with the experimental procedures adapted from a classification test (Graham-Ernhart Block Sort Test, 1963) GEBST.

Review of Literature

Theoretical Aspects of the Problem

"A logical class is a union of elements possessing a common quality. Adding consists in defining the smallest class that contains both" (Piaget, 1952). Classification ability does not suddenly emerge but its successive beginnings can be traced as growing out of earlier abilities. Between 6 to 8 and 18 to 24 months, Inhelder & Piaget(1964) note a child when given a familiar object recognizes its possible uses. These authors suggest that this is a sort of practical classification, however far removed from the operational structure.

Inhelder & Piaget(1964) described three main phases in the development of free classification. The first phase, "graphic collections,"
occurred when the child was unable to separate the properties of the object he grouped from the arrangement he produced. An example of graphic collections would be the arrangement called a "house" by a child which was created by a square and a triangle put on top for a roof. Substages existed in the "graphic collections" but these stages could overlap and do not always appear in the same sequence. The second phase in the development of free classification was "non-graphic collections" in which objects were grouped according to similarity alone. An example of this would be when the child placed all the squares together and placed all the triangles together separated from the squares. Again substages exist, with the final stage exemplified by all groupings being based on the same criterion and subdivided according to the same criterion. The third phase was classification in the strict sense. The understanding of "class inclusion" was the last ability to be achieved by children. Class inclusion referred to the understanding of the child that "all of A are some of B."

A descriptive study (Annett, 1959) of the way in which normal subjects ages 5 to 73 years classify familiar objects demonstrated significant age differences. Younger subjects tended to form many groups. With an increase in subject age (9 years or older) the diversity of the groupings was reduced and subjects tended to sort within the four main categories that existed in the study (animals, plants, vehicles and furniture).

Children's ability to order picture sequences (Swayze, 1967) confirmed Piaget's general idea of a movement from perceptual to conceptual thought with an increase in subject age. An investigation
(Whyte, 1970) of classification ability of children below average intelligence confirmed Piaget's results that classification was not one ability, but a number of related ones which organize in the sensorimotor period and grow out of the ability to bring similar objects together. A summary of Piagetian research in child development (Piaget & Inhelder, 1969) stated that it was possible to trace at the preoperational level the successive beginnings of what would become the additive and multiplicative "groupings" of classes and relations. Piaget & Inhelder (1969), note that the child's preconceptual period occurs approximately between two and seven years, and these authors have noted that the component operations of logical classification emerge during this time.

The following investigations, Inhelder & Piaget (1964); Inhelder (1962) and Piaget (1953 & 1954) assign a nonsignificant role to language. It is noted by Inhelder (1962) that the operations constituting logical classes show continuous progress from elementary behaviors to anticipatory processes that precede and go beyond linguistic associations. Two studies (Sigel & Shapiro, 1966; and Vinacke, 1952) note that a child may be able to perform classification tasks, yet be unable to verbalize the principle underlying his groupings.

A study (Bruner, Olver, Greenfield, et al., 1966) explains classification in terms of language, stating that older subjects translate the image into a form that can be coded into language thus formulating a set of verbal rules to guide them in making transformations.

An early work in Russian in 1934 (Vygotsky, 1962) described the intellectual development of the child in terms of a classificatory structure. Vygotsky (1962) described three basic phases in the devel-
opment of classification. The first phase was called "heaps of objects" and a number of seemingly unrelated objects were placed together. The second phase was called "complexes" and the child began to group blocks on the basic similarities shared by the blocks. This phase was broken down into stages which demonstrated a progressive utilization of similarity until finally the child grouped according to one consistent attribute. The third phase utilized conceptual thinking. Vygotsky (1962) also stressed the necessity of language for activity utilizing all basic intellectual functions.

Agreement exists between Bruner et al. (1966), Inhelder & Piaget (1964), and Vygotsky (1962) on the aspect that concepts develop early and are complete only with increased age and development. Opposing views exist on the importance of language. The researchers, Inhelder and Piaget, state that the precursors of classification start before language development and finally reach a level independent and beyond language ties. In contrast, Bruner et al. (1966), and Vygotsky (1962) state that language is essential in the development of classification performance and in its functioning.

Validation Research on Classification Theories

A statement was made (Flavell, 1963) that there seemed to be no attempt towards validation of Inhelder & Piaget's (1964) work on classification. Since then recent issues of psychological journals (late 1960's and early 1970's) have reported a number of free classification studies. Free classification (Kofsky & Osler, 1967) has become accepted to mean that the subject may choose the criteria of sorting. The term was used in this manner by Kofsky & Osler (1967); Denney (1972a,
1972b); and Overton, Wagner & Dolinsky (1971). It would seem that its usage is accepted. However, Denney (1972a) referred to a number of classification investigations labelling all as free classification research, and some of these so labelled did not fit the definition given previously. Care must therefore be taken in assessing what is meant by the term free classification task.

On reviewing recent classification studies, only two, Denney (1972a, 1972b) attempted to replicate the task of Inhelder & Piaget (1964). Denney (1972a) used two classification procedures, one similar to that employed by Vygotsky (1962) and one similar to that employed by Inhelder & Piaget (1964). Subjects for this research were eight male and eight female children from each of the following groups, 2, 4, 6, 8, 12 and 16. Stimuli were wooden blocks which varied in both color and form. Neither Vygotsky's nor Inhelder & Piaget's results were replicated by the procedures that were meant to correspond to their own. The results did indicate a significant relationship between age and type of response. The same developmental change was observed with each procedure, namely that with increasing age response variability between subjects decreased and the "form" response ultimately predominated. Although this study indicated there might not be developmental stages similar to those reported by Vygotsky (1962) and Inhelder & Piaget (1964) it was suggested that the age range of subjects was too broad for some stages to become apparent. The number of subjects receiving each procedure in each age group was small--eight subjects in an age group received each instruction. To analyse the results Denney (1972a) combined responses into categories. With only eight subjects in each age group, combin-
ing responses into similar categories could eliminate some of the stages reported by the earlier researchers. It is suggested that this procedure could further hinder observation of all the stages previously described.

A significant relationship was found between procedure and type of response. Denney (1972a) stated, "The differences in the types of responses obtained with the two types of procedures was not surprising since Vygotsky and Inhelder & Piaget themselves obtained different results when using different procedures." A larger variety of responses were obtained with the free grouping procedure of Inhelder & Piaget (1964). With the verbal labeling procedure of Vygotsky (1962), however, the subjects were simply instructed to pick out all the "wugs" (or "slibs") and were not instructed to do it any way they wanted. Thus with more restrictive instructions there were fewer types of classification responses.

In a second study, Denney (1972b) attempted to replicate responses obtained by Inhelder & Piaget (1964) and presented children ages 2, 3, and 4 years old with a free classification task similar in nature to the earlier researchers. Many of the responses obtained by Inhelder & Piaget (1964) were not obtained, however the data of Denney (1972b) indicated stages in the development of the ability to group according to similarity. First, a stage occurred in which the child was unable to group according to similarity. Second, there was a stage in which the child grouped only partially according to similarity, and in the third and final stage the child grouped the entire stimulus array according to similarity. The stage labelled "graphic responses" by Inhelder & Piaget (1964) did not
appear as a separate stage in the development of classification in the results of Denney (1972b). In summary, the gradual development of classification ability with an increase in age as reported by Inhelder & Piaget (1964), and Vygotsky (1962) was confirmed by the research of Denney (1972a, 1972b).

None of the other free classification studies attempted to replicate the work of Inhelder & Piaget (1964) but instead focused on a particular aspect of their research. In two studies (Croll, 1970; and Feldman, 1969) children matched geometric shapes that "looked alike." Croll (1970) reported a classification task had methodological advantages for the study of generalizations in young children. With finely graduated changes in the stimuli, it was possible, the author suggested, to accurately obtain estimates of the child's generalization of "circle" for example. A block sorting task by Feldman (1969) investigated the importance of negation by requiring subjects to sort geometric stimuli which were "not X." The results of Feldman (1969) supported Inhelder & Piaget (1964) in the order and approximate age children developed an understanding of negation.

There still appears to be large areas of the theory presented by Inhelder & Piaget (1964) that have not been investigated and only one researcher, Denney (1972a, 1972b) has attempted to replicate the stages in the development of classification performance.

**Conceptual Functioning Assessment and Classification Tasks**

Some earlier researchers (Bolles, 1937; Goldstein & Sheerer, 1941;
Hanfmann & Kasanin, 1937; and Reichard, Schneider & Rappaport, 1944) were interested in free classification as a tool to assess "mental deviation." Distinctions were made between classification on a "concrete" and "abstract" level. The definition of these terms (Vinacke, 1952; and Werner, 1948) was the following: concrete behavior was a response to separate objects, and abstract behavior was a response to the object apart from the perceptual field, and with resultant relations grasped. Goldstein & Scheerer (1941) made distinctions between the two modes of behavior (abstract and concrete) and noted a normal adult utilizes both behaviors whereas an abnormal person has only concrete behaviors. These researchers state that young children use concrete behavior and do not have class concepts like adults. One author (Hunt, 1961) suggested that perhaps the distinction between abstract and concrete thought was but a vague formulation of the same distinction Piaget had made between formal and concrete operations.

The free classification test GEBST, selected as a suitable instrument to assess conceptual development in this research, was used as the "conceptual test" in a test battery to assess brain damage. The research which used the GEBST (Ernhart, Graham, Eichman, Marshall & Thurston, 1963; Graham, Ernhart, Thurston & Craft, 1962; and Graham, Ernhart, Craft & Berman, 1963) demonstrated that conceptual ability was the single most impaired function in the studies with brain damaged children. In devising the test, Graham et al. (1963) adapted procedures from preschool level tests that had been used successfully with older groups to differentiate between brain injured and non brain injured. The GEBST was described by Graham et al. (1963).
The test "required the subjects to match or group blocks which differed with respect to three dimensions; color, size and form. There were four levels of difficulty. Level I involved placing blocks in form boards; Level II required matching of blocks; and Level III and IV required sorting. On Level III the blocks differed on one dimension and were constant in respect to the remaining dimensions. On Level IV blocks differed on two dimensions simultaneously while the third dimension was held constant. At each level all three concepts were tested in separate trials. The procedure is similar to that employed by other tests of the ability to group according to concepts, Goldstein & Scheerer (1941)." The GEBST was standardized on a group of 245 normal and 70 brain injured children ranging in age from 2 years to 5½ years. The general trend on the GEBST was towards negative skewness and decreasing variance with increasing age. Single session reliability was difficult to obtain since this test had too little replication of items of comparable difficulty or items similar in kind. The test-retest reliability coefficient for the GEBST was judged reasonably satisfactory (r=.61). It was stated that in most research, consistency is generally less in a preschool group than with older subjects and the brief procedures were repeated after an unusually long interval, six months. Age was the most important variable affecting performance on the test. Sex and socioeconomic status were less important, but in general girls were superior to boys, and the performance of status groups was in the socioeconomic order of high status groups receiving high performance scores and low status groups low performance scores. Validity was discussed in the research of Graham et al. (1962) and the GEBST was described as the single most discrimin-
ating test in the battery used to distinguish brain injured from normal groups. The three studies cited, Ernhart et al. (1963), Graham et al. (1962, 1963) are all the published research that has used the GEBST. The test manual for the GEBST is in Appendix A.

The GEBST was selected for this particular investigation for several reasons; it purported to assess conceptual ability; it was a free classification test and it was simple to administer. The test could easily be adapted to the aspect of dimensionality without major procedural alterations, it presented a task similar to Inhelder & Piaget (1964), and the scoring system allowed statistical analysis. It was realized that the purpose for which the test was developed was not the same as it was to be used in this study, however, the test authors reported the test measured conceptual ability. It should therefore measure conceptual ability in a sample of exclusively "normal" subjects. The age range of subjects upon which the test was standardized was less than the ages to be studied in this investigation. It was felt that presenting pictorial representation of the geometric shapes would make a more difficult test for the older subjects. The only levels utilized for this research were Level III and Level IV, and additional scoring procedures were carried out to increase the spread of classification performance at the top of the scoring scale in Level IV. The final reason for selecting the GEBST was the possibility it presented for comparing the results of this investigation with other research. There have been many investigations into young children's classification performance, but each investigator developed his own classification task. Comparison of results was thus difficult and possibly could be confounded by the type of classification task used. If the GEBST is a
suitable instrument to assess classification performance, then possibly it could be used by other researchers in this area.

Training Procedures and Classification Research

Many classification investigations attempted to verify or disprove certain aspects of Inhelder & Piaget's (1964) developmental theory of logical thinking. Piaget described, "The central mechanism of intelligence is found in the construction of operations which derive from the general coordinations of actions" Furth (1969). Logical operations, Inhelder & Piaget (1964), are closely linked with certain elementary actions, for example putting things in piles. The development is continuous; after the actions there are various adjustments to these actions, and these in turn become increasingly complex. Maturation, learning, and social education, including language, play a part in the development of logical operations. Logical operations, however, would not occur without the process of the child's actions on the changing environment. Induction of a concept without adequate cognitive organization results in a concept limited to the training or similar situation. Broad concept generalizations can not be performed.

Research with kindergarten children (Aldrich, 1970; and Robinson, 1970) supported Piaget's proposal that training in logical thinking was not possible until the child was ready for this operation. Aldrich (1970) attempted to teach class inclusion and reported an "isolated schema" successful only with the training materials. In his study, Robinson (1970) attempted to develop a variety of conditions to facilitate classification in kindergarten children but
"conditions" did not make a significant difference in the subjects' performance.

Some research (Jacobs & Vandeventer, 1971; Nowak, 1969; Overton & Brodzinsky, 1972; Parker, Rieff & Sperr, 1971; and Youniss, 1971) reported significant differences with training in their subjects, ages six and older. Piaget suggested that the pre-operational stage occurs in children ages 2 to 7 years. The researchers who reported a change from pre-operational classification to concrete operations with training in their older subjects may have provided the environment upon which the child, at that maturational level, could utilize previously developed actions and thus demonstrate true classificatory behavior. The children were in the age group in which Piaget suggested concrete behavior emerged.

One study (Olmsted, Parks & Rickel, 1970) trained lower class negro kindergarten children, middle class five year olds, and gifted four year olds. Significant changes in cognitive skills were reported. These researchers reported changed performance with training material. Piaget however, would question the existence of changes in cognitive level when applicability to new situations was not investigated.

In summary, conflict exists in the reported success of "training subjects in the logical operations of classification." It is suggested that the explanation for the conflict lies in the researchers' assessment of changed cognitive skills. Some researchers reported changed cognitive skills with the test material alone whereas others tested with new stimuli varying in degrees of similarity with the original test materials.
Subject Characteristics and Classification Research

All research reviewed indicated a steady improvement in classification performance with an increase in age of the preconceptual child. Graham et al. (1963) reported age to be the most important variable affecting performance on the GEBST. When attempting to replicate Inhelder & Piaget's (1964) research on classification, Denney (1972b) reported a significant relationship between age and type of response. An increase in the frequency of true classification responses occurred with an increase in subject age from 2 to 4 years. Annett (1959) presented a classification task to a wide age range of subjects. It was reported that sortings showed a clear change with the number of categories reduced as age increased. Explanations for the categorization were recorded by Annett (1959) and the ordinal scale of rating responses demonstrated abstract basis of sortings increased with subject age to adulthood. Regardless of the procedure, type of subject or stimulus object there were no contradictory statements concerning the relationship of age and performance.

Research has been conducted in other cultures to verify some aspects of Piaget's theories. A review of the research in other cultures (Dasen, 1972) noted that most of this research has been conducted at the preoperational and concrete operational stages. The age at which illiterate, primitive subjects reached a designated stage in Piaget's developmental sequence differed from the age range reported by Piaget. The order of the developmental stages was reported by Dasen (1972) to be the same as described by Piaget. Price-Williams (1962) compared classification of literate and illa-
literate 6 to 11 year olds from a primitive African tribe using indigenous material. The African subjects reached concrete operations at an older age (sometimes into adulthood) than Piaget's subjects. All, however, reached this stage.

Investigations have been done comparing classification behavior in different socioeconomic groups. Four investigations (Overton, Wagner & Dolinsky, 1971; Sigel & Shapiro, 1966; and Wei, 1967) compared classification performance in lower and middle class children ranging in age from 3 to 9 years. The study by Sigel & Shapiro (1966) reported children from lower socioeconomic homes scored lower than children from middle class homes when classifying pictorial representation of objects. When the stimulus was actual objects there was little difference between the children from different socioeconomic backgrounds. A study by Wei (1967) reported culturally disadvantaged children scored lower on classification tasks than children from middle class homes. Wei (1967) presented subjects with four classification tasks developed from the work of Inhelder & Piaget (1964). Reporting similar findings to the previously cited research, Overton et al. (1971) explained the lower performance scores of the lower socioeconomic group in this manner. All children may develop the necessary cognitive structures at a similar rate, but activation for lower class children may occur later as the environment fails to provide sufficient opportunity to develop the cognitive structures. This explanation is in harmony with Piaget's theory of development of logical structures by the actions of the child with his environment.

In summary, it appears socioeconomic status is related to
classification performance and the lack of opportunities available to practice classification techniques by the lower socioeconomic children could explain the differences in performance scores between different economic status groups.

Sex differences in classification tasks are: 1. not analysed or discussed, 2. reported to be non-existent or 3. reported to be slight with girls performance scores slightly ahead of boys. Denney (1972a, 1972b) reported no sex differences in free classification performance. A study to standardize a test battery by Graham et al. (1963) reported girls scored slightly higher than boys on the GEBST. It would appear from the literature reviewed, that sex differences on classification tasks are slight, if they exist at all.

Stimulus Attributes and Classification Research

The type of objects researchers use in their investigations affects classification performance. Material which was familiar to the subjects was classified more abstractly than non-meaningful material (Asch, 1970). Disadvantaged preschool children were able to classify pictures abstractly. Using a training procedure in classification, Asch (1970) reported that more abstract classification occurred with the meaningful representational pictures than with non-meaningful geometric forms. A study by Feldman (1969) to assess children's understanding of negation as a logical operation in a classification task reported that more correct responses were made with familiar material (for example geometric shapes) than with unfamiliar material (for example random shapes). Price-Williams (1962) investigated abstract and concrete modes of classification in a
primitive society. This researcher reported that the use of Western material such as geometric shapes did not provide a stimulus object familiar to the illiterate subjects and thus subjects could not group abstractly. A comparison of classification performance of literate and illiterate African subjects on indigenous material resulted in very little performance difference. Wieman and Guthrie (1972) attempted to assess the influence of age and the cultural familiarity of stimulus items on children's categorizing responses. There were two objectives in this research: first, measurement of conceptual growth across ages, and second, a comparison of classification performance with two sets of stimulus items matched for type, directions and difficulty but differing in cultural familiarity. Items from a Philippine non-verbal intelligence test were matched with objects which were culturally familiar. Subjects were age 4 through 9 with 15 male and 15 female in each age group. Results of this study demonstrated non-conceptual sorting explanations showed a decrease with age. A consistent increase in responses interpreted as signifying conceptualization was noted with an increase in age. In dealing with the Philippine items subjects at all ages were less successful in utilizing conceptual strategies. In summary a more abstract classification performance can be obtained from subjects if the stimulus objects are familiar to them.

Some research into classification performance, Annett (1959); Inhelder & Piaget (1964); Sigel (1954); and Sigel & Shapiro (1966), presented tasks with stimulus objects which were normally present in the surroundings of the subject. Annett (1959) used line drawings of objects such as animals (cow, bird), plants (tree, flower), and furn-
iture (chair, television). Inhelder & Piaget (1964) used flowers, pictures of animals, and small toys as stimulus objects. A study by Sigel (1954) compared the classification performance using pictures, word names of meaningful objects and three dimensional toy representation of the objects. There was no significant differences regardless of stimulus presentation as the child responded in a similar fashion to toys, pictures or word names. The conceptualization used by the subjects was based primarily on the meaning of the object. Concepts related to the stimulus properties, such as color, texture, size or form tended to be ignored. In summarizing the research with familiar objects, the categorization of such objects is greatly affected by the particular meaning the subject has for each object. A conclusion by Sigel (1954) was that the meaning of an object was not only independent of the stimulus characteristics of the object, but was also apparently dominant as a basis of organization of the objects. "Meaning dominance" was the term developed by Sigel (1954) to explain this result.

Dimensionality of stimuli has been demonstrated to affect performance of young children on a variety of tasks. Dimensionality refers to the form of the stimulus object. Stimuli can be three dimensional objects such as balls, blocks or small toys. Two dimensional objects are flat cut outs of objects such as circles, squares, form board or puzzle pieces. Pictorial representation, or line drawing on cards is the third type of stimulus dimensionality that can be utilized.

Sigel & Shapiro (1966) used stimulus material consisting of two dimensional pictorial representation of familiar objects and actual
life sized familiar objects. The categorization behavior of lower
and middle class negro preschool children was compared with the two
types of stimuli. Middle class subjects were able to perform in a
similar manner with both types of presentation, however the lower
class subjects were confused by the pictorial representation even
though they could identify the object. According to Asch (1970)
"disadvantaged" preschool children were able to classify pictures
abstractly. This is contradictory to the results of Sigel &
could be resolved by attributing the success of Asch's subjects to
training they had received. Stimulus dimensionality according to
the research of Sigel & Shapiro (1966) could be an important factor
affecting results which compare performance of preschool children
from different socioeconomic levels. The research of Sigel &
Shapiro (1966) was the only classification study reviewed which in-
vestigated the affects of stimulus dimensionality.

Many researchers have investigated the relationship of stimulus
dimensionality and discrimination performance. Studies in this area
(Etaugh & Van Sickle, 1971; Everett & Armstrong, 1968; Falk, 1968;
and Stevenson & McBee, 1958) presented subjects with a discrimination
task and all reported that three dimensional stimuli were discriminated
more rapidly than pictorial representation or line drawings of the
same objects. The age range of subjects in the four cited studies
was 3 to 6 years. The age of the subject was reported to be related
to the child's ability to discriminate pictorial material as the
youngest subjects were less able to discriminate pictorial compared
to three dimensional stimuli. A study (Dornbush & Winnick, 1966)
reported five year old subjects could learn a discrimination task more rapidly with three dimensional stimuli than with pattern stimuli. No significant differences in the discrimination learning time occurred between three dimensional and pattern stimuli with seven year old subjects. A similar result was reported in another study by Falk (1968). The conclusions of this study were that the rate of learning a size discrimination was affected by the dimensionality of the stimuli and this effect was greater for the four year old subjects and less for the six year old subjects. Summarizing the research relating stimulus dimensionality and age, it appears that for children under six years of age, pictorial material is more difficult to discriminate than three dimensional objects. Around six or seven years, however, the child becomes able to discriminate both pictorial and three dimensional material in a similar manner.

Some researchers, Engle (1972); Etaugh & Van Sickle (1971); Gibson (1963); and Falk (1968) attempt to explain the effect of stimulus dimensionality in a discrimination task. Two explanations are suggested. The third dimension, or "depth cue," gives additional information that is lacking in pictorial representation and thus enables subjects to discriminate more rapidly with three dimensional than pictorial stimuli. The second explanation suggests the importance of the information a subject gains through touching the three dimensional objects in contrast to the visual action alone on pictures. This second explanation was supported by two studies, Engle (1972) and Tanaka (1968), which compared the effects of different instructional methods on classification performance. Greater gains in classification performance scores resulted when subjects were
able to touch stimulus objects than when they visually looked and verbalized without touching stimuli.

Contrary to all the previously cited research on pictorial presentation, data was presented (Devor & Stern, 1970) demonstrating that three dimensional stimuli do not result in an easier task than the use of pictures of the same object. The reason for the different results of Devor & Stern (1970) could be the type of task these researchers presented, and the scoring procedures for subject performance. Verbal labels were taught and then rote recall of the correct label was required. The learning of the labels was not facilitated by the three dimensional objects. The authors noted subjects played with the objects and were thus distracted from the verbal taped instructions. Previously cited researchers, Inhelder & Piaget (1964); Piaget (1953, 1954); Sigel & Shapiro (1966); and Vinacke (1952), explained that verbal ability was not the same as conceptual knowledge in young children. Devor & Stern (1970) appeared to be scoring verbal ability and labelling it conceptual performance. In summary, stimulus dimensionality was a factor which affected performance of young children on discrimination tasks.

The relationship of the number of objects to classification performance was investigated. A study by Feldman (1969) of children's understanding of negation as a logical operation in a classification task reported that the number of objects in the task interacted with subject age. Reducing the number of stimulus objects did not help the three year old subjects, as this age was still unable to handle negation, nor did it help the seven year old subjects who had a full understanding of the task and could successfully perform the task.
with the complete stimulus set. The performance of four and five year old subjects was improved when fewer objects were presented. Feldman (1969) was the only researcher with a classification task who varied the number of stimulus objects to assess the effect of this variable on performance scores.

The perception of color and form by young children has been investigated by many researchers. A review of this area by Sigel (1954) cited research that demonstrated both form and color discriminations can be made by very young infants. There is conflicting research evidence concerning color versus form preference by subjects three and four years of age. In a color-form sorting task (Harris, Schaller & Mitter, 1970) form matches were more frequent for geometric figures. There was a significant increase in the number of form matches with an increase in subject age. The subject age in this investigation was preschool to grade three students. Kagan & Lemkin (1961) used a task of matching a stimulus object to a standard to study form, color and size in children's conceptual behavior. Children, ages three to eight years, preferred to match by form rather than color as a basis of similarity. Researchers (Modreski & Goss, 1969) instructed young children to give the name of colored geometric shapes. The results were similar to Denney (1972a) and Suchman & Trabasso (1966). Young children selected form names more often than color. Even three year old subjects named by form more often than color, and the number of subjects naming by form increased with ages four and five.

Other researchers reported results which conflicted with those just cited. A recent study, (Denney, 1972a) reported an increase in
the use of color as a basis of sorting up to the age of six, and then a decrease. A study (Suchman & Trabasso, 1966) which required children to "point to the two alike" from slide presentations demonstrated that 3 year old children tend to match by color and 6 year old children by form. The children ranged in age from 2 years 10 months to 6 years 6 months in this investigation. A study which used a discrimination task to assess young children's attribute preference reported that color was preferred more than form with nursery school and kindergarten subjects (Trabasso, Stave & Eichberg, 1969). In the color versus form conflict just presented, color has not been clearly established as the preferred dimension by which to match objects when subjects are young children. One researcher (Corah, 1964, 1966) presented conflicting evidence in his two studies. Corah (1964) reported color to be clearly the dominant characteristic, and then later Corah (1966) stated that form was preferred to color even by the youngest subjects. He presented a similar match to standard task each time. Preschool children were required to match geometric forms, line drawings of a simple form, and an outline of asymmetric forms. Corah could not adequately explain the conflicting results. In summary, there are more studies supporting color preferences than form preferences in the preschool children, but it has not been clearly demonstrated that color is preferred by the majority in this age group.

Research in the area of color versus form preference by school age subjects presents consistent findings. All the previously cited research on the color versus form preference stated that their older school-age subjects matched or categorized by form in preference to color.
In the research of Kagan & Lemkin (1961) and Sigel (1964) children's performance with three stimulus characteristics (form, color and size) was compared. Size was reported to be a less salient cue than either form or color. Sigel (1964) stated that matching or classifying by size required ignoring the more observable form and color. An object could be one particular color or form but an additional step of relating two or more objects was required when size was selected as a basis of categorization. Size existed only in relation to another object, and thus a comparison was required before a matching or a classification task could be performed.

Procedures in Classification Research

In a review of research methods in child development (Bijou & Baer, 1960) the point is raised that a special problem occurring in research with young children is: "Does the child understand what is required of him?" The studies of classification by Inhelder & Piaget (1964), used the instructions, "Put together things that are alike." Other instructions similar in nature were also used by these researchers but it did not appear to affect their results. Two instruction wordings were compared by Annett (1959). With a mean subject age of 8 years there was no significant difference in performance when the two instructions were used for the classification tasks. Annett (1959) first used the instructions, "Sort them into groups so that the ones that go together are in the same group." This was compared with the instructions, "Put those that belong together into groups." An early classification study by Bolles (1937) instructed subjects to, "Put those together that belong together." In a recent
study, Denney (1972b) instructed preschool children to, "Put the things that are alike, or the things that go together, into groups."

In another free classification study, Kofsky & Osler (1967) instructed young children, "Put together what looks the same." In a review of Piaget's experimental methods (Oleron, Piaget, Inhelder & Greco, 1969) which discussed how the instructions vary according to the aim of the experiment, it was noted that "indeterminate instructions," such as the ones just previously quoted enable the experimenter to observe spontaneous tendencies, for example the way the child would independently group objects. Instructions that are not restrictive, according to these authors, Oleron et al. (1969), will maximize classification performance, or in other words, result in the highest classification performance scores.

Some researchers presenting matching to sample tasks used similar instructions to the ones previously discussed. Bearison & Sigel (1968), Harris et al. (1970) and Kagan & Lemkin (1961) used the following type of instructions, "Which one of these is the same or alike as this (standard)?" In these three studies there was no discussion of problems encountered with children unable to understand the task. The subject age range in these investigations was 3 years to 11 years.

In summary, none of the research in free classification and matching sample studies which utilized "indeterminate instructions" reported subjects unable to understand what they were expected to do. One could conclude that this type of instruction was suitable for young children presented a categorizing task.

"Indeterminate instructions" may enable the experimenter to observe "spontaneous tendencies" as suggested by Oleron et al. (1969).
Unfortunately the wide variation in classification performance results in difficult evaluation of individual performances. Inhelder & Piaget (1964) described stages in the development of classification performance by observing many categorizations and noting similarities that occurred between subjects. The performance of an individual subject was not scored, but instead the general behavior pattern across ages was described. Grouping of response patterns and reporting on the frequency of their occurrence was also the evaluation procedure followed by Annett (1959) and Denney (1972a, 1972b).

With more structured tasks, a variety of scoring systems were used to evaluate the conceptual level of performance. Vygotsky (1962) developed an extremely structured and complicated classification test. It was evaluated by counting the number of separate steps or hypotheses tried by a subject to solve the problem. The evaluation of a subject's performance was simplified by the design of this test. The test was far too difficult for preconceptually aged subjects and the restrictive instructions would not allow "spontaneous tendencies." Wei (1967) presented tasks modelled on Inhelder & Piaget (1964) classification tasks. Each subject was asked a standardized form of question with a score of one point for each correct response. An additional scoring procedure was followed by Wei (1967). Analysis of responses to each item within a task was made in terms of percentage of success for the item. This second evaluation procedure was more subjective and would result in difficulties analysing individual performance.

The GEBST has indeterminate instructions and a very similar task to that of Inhelder & Piaget (1964). It is designed with a slight structuring so that individual performance can be scored. The
structuring of the test is not sufficient to allow test scoring free of experimenter judgment. Scoring of a free classification task is still a major problem which can be solved by increasing the task structure. This results in elimination of some of the "spontaneous" classification performances. The complete solution of the scoring problem of free classification tasks has not been demonstrated by any of the reviewed research.
Chapter 2

The Research Problem

Purpose

The purpose of this investigation was to examine the relationships among stimulus object dimensionality, other stimulus characteristics, subject characteristics and preconceptually aged children's performance on a conceptual task.

Operational Definitions

For the purposes of this study the major variables are defined as follows:

**Dimensionality** of the stimulus refers to the physical description of the stimulus object. It denotes whether it is a geometric solid (termed a three dimensional object), a flat geometric shape (termed a two dimensional object), or a picture of a three dimensional geometric solid (termed a pictorial representation).

**Stimulus characteristics** refers to the color, form or size of the stimulus object.

**Classification performance** refers to the score on difficulty Level III and/or Level IV on the GEBST. The scoring on Level III is according to the manual for the GEBST, and on Level IV, the scoring is according to the manual plus additional changes noted under procedures and scoring.
Level III refers to the difficulty level of the GEBST where six stimulus objects are presented which differed on one dimension and were constant on the remaining two.

Level IV refers to the difficulty level of the GEBST where blocks differed on two dimensions simultaneously, while the third dimension was constant.

Trial refers to which of three tasks were presented in both Level III and Level IV of the GEBST.

Subject age was the child's age in years and months at the time of testing.

Hypotheses

Two major hypotheses and three minor ones have been formulated. The major hypotheses refer to the primary interest of this investigation; these are stimulus dimensionality and subject age. The three secondary hypotheses are concerned with the variables secondary in priority in this investigation, such as the other stimulus characteristics and their relationships with subject age and dimensionality.

Major Hypotheses:

1a Dimensionality of the stimulus affects performance in the following manner: classification performance scores are higher with three dimensional and two dimensional objects than with pictorial representation.

2a Subject's age affects classification performance in the following manner: classification performance scores are higher with older subjects and lower with younger subjects.
Minor Hypotheses:

1b Characteristics of the stimulus affects performance in the following manner: classification performance scores are higher with color and form based sorts than with size based sorts.

2b Stimulus size and stimulus dimensionality interaction affects performance scores in the following manner: size based classification performance scores are higher with two and three dimensional stimuli and lower with pictorial representation.

3b Subject age and stimulus characteristic interaction affects classification performance scores in the following manner: size based classifications result in lower scores than color or form for the youngest subjects, but with an increase in subject age there is a smaller difference in performance scores between size, form and color based sorts.
Chapter 3

Methods

Subjects

The 90 subjects for this investigation were from the following five age groupings with 18 in each age group: preschool 3 and 4 year olds, kindergarten, grade one and grade two children. All subjects resided in a middle class area in the city of Vancouver. Twelve of the 3 year olds and twelve of the 4 year olds were attending a cooperative preschool. The remaining six subjects for each of these groupings were attending a cooperative daycare center at the University of British Columbia. Kindergarten, grade one and two subjects attended a public elementary school in Vancouver. Approximately equal numbers of male and female children were subjects. The distribution of sexes is summarized in Table 1. Subjects were selected by age group, and equal numbers of each sex was not a criteria in subject selection as there were insufficient subjects to do this in the 3 year group. It was not attempted for any other age group.

The age range of the subjects varied within each grouping. To obtain 18 subjects 3 years of age, all 3 year old children registered at both centers were participants in this research. In the 4 year group all the names of children registered at a center were placed in a box. Twelve names were drawn from the preschool 4 year group and six from the daycare group. In order to obtain similar subjects for each age group the same proportion of children from each center
Table 1

Distribution of Subjects by Age and Sex

<table>
<thead>
<tr>
<th></th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>9</td>
<td>8</td>
<td>41</td>
</tr>
<tr>
<td>Male</td>
<td>11</td>
<td>10</td>
<td>9</td>
<td>9</td>
<td>10</td>
<td>49</td>
</tr>
</tbody>
</table>
participated in the study. At the elementary school, 30 arbitrarily selected children in each age group were given a letter to take home explaining the test procedure. Consent from a parent for all individual testing of children in the Vancouver public school system was required. At the bottom of the mimeographed letter explaining the research was a section for the parent to tear off, sign, and return to the school. The letter that was sent is in Appendix B. Only children who returned signed consent forms could be subjects. More than the required number of consents were returned and 18 subjects were arbitrarily selected from each of these groupings; again by placing names in a box and drawing the required number.

Subject placement in the grouping for testing was the same as school placement; for example, children in the preschool 3 year old group were grouped as 3 year olds even though by the time of testing in February, a few of the group had January birthdays and were actually 4 years 0 months. All children were in the normal grade placement, and the only "apparent" misplacements were children who had January birthdays. Each age grouping, however, contained a typical spread of ages.

**Stimuli**

There were three sets of stimulus objects; three dimensional styrofoam objects, two dimensional flat styrofoam objects and pictures of the three dimensional flat styrofoam objects. Pictorial representation was accomplished by photographing the three dimensional styrofoam objects. Each stimulus set contained objects which varied in the following ways; color (red, yellow and blue), form
(eg. triangle, circle and square), and size (large, medium and small).

The dimensions of the three dimensional and flat large objects were 3 inch square, 3 inch triangle side, and 3 inch circle diameter. The relevant dimension for each form in the medium size was 2 inches and in the small size 1 inch. The flat objects were cut from 1/2 inch styrofoam. The photographs of the three dimensional objects were 3 5/8 inches by 4 5/8 inches mounted on cardboard and sealed in thin, clear plastic. Appendix C contains three illustrations of the pictorial representation scaled down by 1/3.

The colors used were strong primary hues of red, yellow and blue. The background for the photographs was grey and the angle of the picture displayed the third dimension of the objects photographed. The size relationships among the objects pictured was retained. There were 26 objects in each stimulus set. There were three stimulus sets; one set for each level of dimensionality. Table 2 lists all the objects contained in a stimulus set.

Procedure

The instruction manual for the GEBST is in Appendix A. This manual was followed except for scoring changes noted in the next section and changes in trial orders discussed in this section. Two difficulty levels were presented to each subject. Level III was less difficult and was presented first. It consisted of grouping objects on one stimulus characteristic with the other two held constant. There were three trials in Level III, one for each stimulus characteristic. Level IV tasks were more difficult and were thus presented after Level III. Level IV objects differed on two stimulus
Table 2
Stimulus Objects

<table>
<thead>
<tr>
<th>Circles</th>
<th>Squares</th>
<th>Triangles</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 large yellow</td>
<td>2 large yellow</td>
<td>2 large yellow</td>
</tr>
<tr>
<td>2 large red</td>
<td>1 large red</td>
<td>1 large red</td>
</tr>
<tr>
<td>2 large blue</td>
<td>1 large blue</td>
<td>1 large blue</td>
</tr>
<tr>
<td>2 medium yellow</td>
<td>1 medium yellow</td>
<td>1 medium yellow</td>
</tr>
<tr>
<td>1 medium red</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 medium blue</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 small yellow</td>
<td>1 small yellow</td>
<td>1 small yellow</td>
</tr>
<tr>
<td>1 small red</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 small blue</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
characteristics with the remaining characteristic held constant. There were again three trials so that each of the three stimulus characteristics were held constant on one trial. The order of trials in the manual was not followed but trial orders were arbitrarily selected in Level III and again in Level IV to obtain 18 combinations of trials. Each child in an age group received a different trial order. Dimensionality of test material was arbitrarily assigned to the 18 selected trial orders within an age group, with six subjects receiving the test in three dimensional stimuli, six with two dimensional, and six with pictorial stimuli. The same combination of trials and stimulus dimensionality was replicated in each of the five age groups. A subject was tested once, and thus each subject performed under just one dimensionality.

Each subject was tested individually, and all tests were administered by the same experimenter. The testing was done in a separate room from the school classroom and the experimenter sat at a table opposite the child. All stimuli were placed on a heavy felt cloth. The felt held the styrofoam objects in the exact position the child placed them. Identical procedure was followed for all groups, except at the daycare center where the experimenter and subject sat on the floor on pillows and the "game" was played on a mat with the same cloth on top. The experimenter sat opposite the child, both on the floor. When using the table the stimulus materials, when not in use, were kept out of sight of the subject by placing them on chairs below the table level. When testing at the daycare center the material was out of sight behind the experimenter.

The subjects classification arrangements on Level IV were dia-
grammed immediately after the child completed a trial. Level III was recorded by scoring as it was not complicated. The subjects were told the experimenter was at school studying about children and wanted to know how children played this game. "Put together the ones that go together," were the instructions given the subject. When questioned if this was the way to do it, or how to do it, the experimenter replied, "Put them together the way you think they belong." No further help was given, and encouragement consisted of thanking subjects after each trial. Between Level III and Level IV the experimenter said that the next part of the game was a little bit different and they would have to look closely. When the experimenter administered the GEBST prior to this investigation it was observed that a few subjects attempted to pair objects on Level IV. This was the correct strategy on Level III and the continuation of this strategy interfered with performance on Level IV. It was hoped this problem would be avoided by warning subjects when they started Level IV that the task was a little different.

Scoring

Each subject's grouping was scored according to the rules of the GEBST with the following exceptions. Level II was completely omitted as this level was too easy for the age range of subjects tested in this investigation. Level III was scored exactly as in the manual; there was a maximum of two points on each trial to give a maximum of six points that could be obtained over three trials. There were changes in the scoring on Level IV. The manual was followed up to and including a score of 9, which was a correct sort on one concept.
A score of 10 was assigned to a correct sort on one concept and one error on a two concept sort. A score of 11 was assigned to a correct sort on two concepts, but the interrelationships of the two concepts not comprehended. A score of 12 was assigned to a correct sort on two concepts utilizing the interrelationships of all objects. There was a maximum of 12 points on each trial. Thus a maximum of 36 points on Level IV could be obtained over three trials.

The final change made in the area of scoring was concerned with the total scores. The manual combined the scores of Level II, Level III and Level IV to produce a total score for each subject. The maximum total score on Level III was six and on Level IV 36 points. These were two different scoring scales and the effects that occurred on Level IV would have numerically overwhelmed any effects that could have occurred on Level III if the scores were combined. The trials on Level III were qualitatively different from those on Level IV. Level III presented all subjects with an identical task resulting in a clear measure of characteristics. Level IV was a choice paradigm and the effects of this choice confounded the characteristic measure. It was for these two reasons that each subject's performance was scored separately on Level III and Level IV.

Design

There were four sources of variance which determined the design of this study: stimulus factors were dimensionality and characteristics, subject factors were age and individual subject differences. For the purpose of the data analysis, subjects were nested under
age and stimulus dimensionality and crossed with stimulus characteristics. This arrangement is characterized in Figure 1. The same design characterizes both Level III and Level IV. Two separate analyses were necessary for Level III and Level IV. The reasons for the two analyses have previously been discussed in the "scoring" section. The same subjects performed the Level III task and the Level IV task. In the diagram on Level III, characteristics were color, form, and size. On Level IV there was a choice paradigm for characteristics. The stimulus characteristic symbols in Figure 1 for Level IV refer to the characteristic not present in the choice. A size-form choice is symbolized by color as color is held constant. A color-size choice is symbolized by form as form is held constant and a color-form choice is symbolized by size as size is held constant. Sex of subjects was of secondary interest and was analysed in an extension of this analysis rather than analysis of variance.

Statistical Analysis

Analysis of variance was used to test the significance of stimulus dimensionality, stimulus characteristic, subject age, and the interactions of these factors. Age was treated as a categorical variable in this analysis. All subjects in one group were thus classified identically in age. Five distinct age groups were thus created by grouping the subjects according to placement in their preschool or school class. Stimulus dimensionality was a treatment variable, and subjects were arbitrarily assigned to one of the three treatment groups; three dimensional, two dimensional or pictorial
<table>
<thead>
<tr>
<th>Age</th>
<th>Stimuli Characteristics</th>
<th>Stimuli Dimensionality</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>Color</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Form</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Size</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Pictorial</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2 Dimensional</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 Dimensional</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cell Size</td>
<td>n=</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Color</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Form</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Size</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Color</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Form</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Size</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Color</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Form</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Size</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Color</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Form</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Size</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fig. 1. Diagram of the Research Design.
representation. Each subject received only one type of stimulus dimensionality. Stimulus characteristic (color, form and size) was also a treatment variable, but all subjects received all three types of stimulus characteristics. A subject thus received all three characteristics but only one dimensionality. Modified classification performance scores on the GEBST was the dependent variable in this investigation.

The effects of age treated as a continuous variable, and sex and the interactions of these two variables, were analysed with least squares regression analysis. Recent studies, (Cohen, 1968; and Overall, 1969) have noted that conventional analysis of variance can be accomplished by least squares regression methods. No attempt was made to control cell frequency in regard to the sex of subject and unequal numbers of the sexes occurred in some cells. The effects of subjects' sex on classification performance could not be analysed with the analysis of variance procedures followed for the other variables as the analysis of variance program required equal cell size. The analysis of variance required the continuous variable of age to be arbitrarily levelled into five age groupings. In the regression analysis the effect of age as a continuous variable was studied.

In his study, Overall(1969) described three methods of regression analysis. The third method was selected to test the significance of the variable sex, the continuous variable age, and their interactions. It was a step-down analysis involving an à priori ordering of the effects. Estimates of each effect were adjusted for those preceding it in the ordering and ignoring those following it.

The strategy for ordering the independent variables of this study
was adapted from Overall (1969). The organismic variables were tested first, then the experimental factors, organismic interactions, experimental interactions and finally all organismic and experimental interactions. The regression analysis was performed separately for each of the two levels of the GEBST. A total of 35 variables were entered in the analysis, though interest centered on only the ones which had not been previously tested with the analysis of variance.
Chapter 4

Results

Two separate analyses of variance were performed. The same analysis model was used for Level III of the GEBST and then repeated for Level IV. In the ANOVA table, "dimensionality" (D) refers to the variable of three dimensional stimuli, two dimensional stimuli and pictorial representation of the three dimensional stimuli. The term "characteristic" (C) refers to the variable color, form and size. Age (A) is the variable of the five arbitrary age groupings of subjects tested. Each subject was only one age and received only one dimensionality, thus individuals were nested under AD, but did receive all three characteristic levels.

The summary table of Level III analysis of variance, Table J, indicates that there was a statistically significant effect of age on classification performance. As a result of analysis of variance, and graphing the means, it may be seen that Level III was most difficult for 3 year old subjects, less difficult for 4 year olds and a ceiling effect then occurred with subjects age 5, 6 and 7 receiving perfect scores.

A statistically significant characteristic by dimensionality interaction effect occurred for Level III. A graph of means for this interaction is illustrated in Figure 3. It may be seen that subject performance did not vary significantly on dimensionality except when sorting pictures. It was significantly more difficult to sort pictures by size than by form or color.
Table 3

Summary of Analysis of Variance of Classification Performance on Level III

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (A)</td>
<td>4</td>
<td>1.1333</td>
<td>5.4643**</td>
</tr>
<tr>
<td>Dimensionality (D)</td>
<td>2</td>
<td>.1778</td>
<td>.8571</td>
</tr>
<tr>
<td>Characteristics (C)</td>
<td>2</td>
<td>.3111</td>
<td>2.6250</td>
</tr>
<tr>
<td>A X D</td>
<td>8</td>
<td>.1778</td>
<td>.8571</td>
</tr>
<tr>
<td>A X C</td>
<td>8</td>
<td>.1445</td>
<td>1.2188</td>
</tr>
<tr>
<td>D X C</td>
<td>4</td>
<td>.3556</td>
<td>3.0000*</td>
</tr>
<tr>
<td>A X D X C</td>
<td>16</td>
<td>.1889</td>
<td>1.5938</td>
</tr>
<tr>
<td>Individuals I(AD)</td>
<td>75</td>
<td>.2074</td>
<td></td>
</tr>
<tr>
<td>IT(AD)</td>
<td>150</td>
<td>.1185</td>
<td></td>
</tr>
</tbody>
</table>

*p < .05

**p < .01
Fig. 2. Mean Classification Performance Scores with the Effects of Age on Level III.
Fig. 3. Mean Classification Performance Scores with the Interaction of Stimulus Dimensionality and Stimulus Characteristics on Level III.
The summary table of Level IV analysis of variance, Table 4, indicates that there was a statistically significant effect of age on Level IV scores. The graph in Figure 4 illustrates the effect of age on classification performance. This graph illustrates that on Level IV the test was most difficult for 3 year old subjects and it became less difficult with an increase in subject age. The task was the least difficult for 7 year olds. There is not a ceiling effect in Level IV as in the case of Level III.

Table 4 indicates that there was a statistically significant effect of characteristics. Figure 5 illustrates the effect of varying the subjects' choice of size, color or form as a basis for sortings. The bar graph in Figure 5 visually portrays that with a choice of form or color based sorts, (size constant) subject mean scores were highest, next were means from a choice of size or color sorts (form constant). Lowest mean scores occurred when subjects had a choice of classification basis from size and form (color constant). The Bonferroni test was used to statistically compare the significance of the difference between the characteristic means illustrated in Figure 5. It may be seen in Table 5 that form or color sorts resulted in significantly different classification scores from size or form sorts.

It is not possible to compare classification performance contrasting size, color and form on Level IV as subjects were given a choice for their basis of classification. Figure 6 was not obtained with analysis of variance but is included in this section because of its relevance to this area. The graph in Figure 6 illustrates which characteristic predominated in the classification performance of
Table 4
Summary of Analysis of Variance of Classification Performance on Level IV

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (A)</td>
<td>4</td>
<td>43.9685</td>
<td>8.7755**</td>
</tr>
<tr>
<td>Dimensionality (D)</td>
<td>2</td>
<td>5.8037</td>
<td>1.1583</td>
</tr>
<tr>
<td>Characteristics (C)</td>
<td>2</td>
<td>8.4704</td>
<td>3.4012*</td>
</tr>
<tr>
<td>A X D</td>
<td>8</td>
<td>4.3407</td>
<td>.8664</td>
</tr>
<tr>
<td>A X C</td>
<td>8</td>
<td>10.8824</td>
<td>4.3698**</td>
</tr>
<tr>
<td>D X C</td>
<td>4</td>
<td>2.2926</td>
<td>.9206</td>
</tr>
<tr>
<td>A X D X C</td>
<td>16</td>
<td>2.7880</td>
<td>1.1195</td>
</tr>
<tr>
<td>Individuals I(AD)</td>
<td>75</td>
<td>5.0104</td>
<td></td>
</tr>
<tr>
<td>IT(AD)</td>
<td>150</td>
<td>2.4904</td>
<td></td>
</tr>
</tbody>
</table>

*p<.05

**p<.01
Fig. 4. Mean Classification Performance Scores with the Effects of Age on Level IV.
Fig. 5. Mean Classification Performance Scores with the Effects of Characteristics on Level IV.
Table 5

Mean Comparison of Stimulus Characteristics on Level IV using the Bonferroni Test with a 95% Confidence Interval

<table>
<thead>
<tr>
<th>Contrast</th>
<th>Means Compared</th>
<th>Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Lower Limits</td>
</tr>
<tr>
<td>1</td>
<td>Size-Color vs. Size-Form</td>
<td>-.1603</td>
</tr>
<tr>
<td>2</td>
<td>Form-Color vs. Size-Form</td>
<td>.0286</td>
</tr>
<tr>
<td>3</td>
<td>Size-Color vs. Form-Color</td>
<td>-.3825</td>
</tr>
</tbody>
</table>

* p < .05
Fig. 6. Preferred or First Selected Characteristics for the Five Ages Tested on Level IV.

a One subject not tabulated as a perfect two-concept sort performed.

b Four subjects not tabulated as all unable to perform this trial.
subjects in Level IV. The total number of the tested subjects making a particular choice is tabulated. Figure 6 reflects choice of characteristics, but does not take into consideration the score subjects received on the classification tasks. The general trend indicated with the size versus color choice was a slight increased use of size with an increase in age, and a corresponding decrease in the selection of color. With a form versus color choice, an increase in the use of form-based sorts and a decreased use of color-based sorts occurred with an increase in age. With a choice of size or form as a basis of classification, at all ages form was most often preferred over size.

Table 4 indicated that statistically different means occur in Level IV with the interaction of age and characteristics. The graph in Figure 7 illustrates this interaction effect. The graph demonstrates similar classification performance with all three characteristic choices except for the youngest subjects. Size or form-based sorts were significantly lower scores with 3 year old subjects but by 4 years of age the same regular pattern of increase in performance with an increase in age occurred for all characteristics. Figure 7 illustrates that the significant interaction effect between stimulus characteristics and age occurred because of the classification performance of the 3 year old subjects. A mean comparison with the Bonferroni test compared 3 year old performance with characteristics. This is shown in Table 6. It can be seen that a size-form sort results in significantly different classification performance than either a size-color sort or a form-color sort. Size-color sorts did not result in significantly different classification performances from form-color sorts.
Fig. 7. Mean Classification Performance Scores with the Interaction of Age and Stimulus Characteristics on Level IV.
Table 6
Mean Comparison of Stimulus Characteristics for the 3 Year Age Group on Level IV using the Bonferroni Test with a 95% Confidence Interval

<table>
<thead>
<tr>
<th>Contrast</th>
<th>Means Compared</th>
<th>Confidence Interval</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Lower Limits</td>
<td>Upper Limits</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Size-Color vs. Size-Form</td>
<td>.0027</td>
<td>5.3303*</td>
</tr>
<tr>
<td>2</td>
<td>Form-Color vs. Size-Form</td>
<td>.3918</td>
<td>5.7194*</td>
</tr>
<tr>
<td>3</td>
<td>Size-Color vs. Form-Color</td>
<td>-3.8916</td>
<td>2.2749</td>
</tr>
</tbody>
</table>

*p < .05
One of the main purposes of this investigation was to study the effect of dimensionality on classification performance. The results were not significant on either Level III or Level IV. To aid in the explanation of the unexpected result, the mean classification performance on Level IV for each dimension is illustrated with a bar graph in Figure 8. This graph illustrates that there were almost identical mean classification performance scores for two of the following types of stimuli; two dimensional, and three dimensional. Pictorial material was expected to be more difficult but the total mean difference was not significant.

On Level III there was a statistically significant characteristic by dimensionality interaction effect, but this effect was not significant on Level IV. To aid in the explanation, the non-significant means for this interaction on Level IV are graphed in Figure 9. Size versus form sorts resulted in slightly lower mean scores on all three dimensionality types. This was not a significant difference such as occurred in pictorial size sorts on Level III. The problem in Level IV, however, was not the same problem as in Level III. A subject was forced to sort by form in Level III and this was demonstrated to be very difficult for some subjects. Level IV allowed a choice of another preferred dimension so the difficulty with pictorial stimuli occurring on Level III could be avoided by a subject.

Least squares regression analysis was used to test the significance of the following: subject age as a continuous variable, sex of subject, the interactions of these variables with each other, and the interactions of these variables with stimulus dimensionality and stimulus characteristics. The purpose of the regression analysis
Fig. 8. Mean Classification Performance Scores with the Effects of Dimensionality on Level IV.
Fig. 9. Mean Classification Performance Scores with the Interaction of Stimulus Dimensionality and Stimulus Characteristics on Level IV.
was to extend findings of the analysis of variance. In Table 7 and 8 only significance tests for the additional terms are summarized though all terms were incorporated in the regression model.

Table 7 indicates the only significant variable in Level III was the age of subject. Subject age produced similar results whether treated categorically in the analysis of variance or as continuous in the regression analysis. Sex was not a significant factor affecting classification performance on Level III, and none of the sex interactions were significant.

Summary Table 8 shows age treated as a continuous variable on Level IV. Age was highly significant with the regression analysis, a result corresponding to that obtained with analysis of variance when age was treated categorically. Table 8 indicates the significance of an age and characteristic interaction, and this corresponds to the finding in the analysis of variance which was graphed in Figure 7 to show the interaction effects. Sex on Level IV was not a significant variable and none of the interactions involving sex were significant.
### Table 7:

Summary of Stepwise Regression of Independent Variable Predicting Classification Performance on Level III

<table>
<thead>
<tr>
<th>Variable</th>
<th>df</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (A)</td>
<td>1/234</td>
<td>19.4879*</td>
</tr>
<tr>
<td>Sex (S)</td>
<td>1/234</td>
<td>.5120</td>
</tr>
<tr>
<td>A X S</td>
<td>1/234</td>
<td>.4819</td>
</tr>
<tr>
<td>A X Dimension (D)</td>
<td>2/234</td>
<td>.9940</td>
</tr>
<tr>
<td>A X Characteristic (C)</td>
<td>2/234</td>
<td>1.8222</td>
</tr>
<tr>
<td>S X D</td>
<td>2/234</td>
<td>1.1747</td>
</tr>
<tr>
<td>S X C</td>
<td>2/234</td>
<td>.2108</td>
</tr>
<tr>
<td>All higher order interactionsa</td>
<td>16/234</td>
<td>1.4096</td>
</tr>
</tbody>
</table>

a All three and four factor interactions were tested as a group and were still not significant

* p < .01
Table 8.
Summary of Stepwise Regression of Independent Variable Predicting Classification Performance on Level IV

<table>
<thead>
<tr>
<th>Variable</th>
<th>df</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (A)</td>
<td>1/234</td>
<td>40.9907*</td>
</tr>
<tr>
<td>Sex (S)</td>
<td>1/234</td>
<td>2.1981</td>
</tr>
<tr>
<td>A X S</td>
<td>1/234</td>
<td>.1238</td>
</tr>
<tr>
<td>A X Dimension (D)</td>
<td>2/234</td>
<td>2.0897</td>
</tr>
<tr>
<td>A X Characteristic (C)</td>
<td>2/234</td>
<td>5.5573*</td>
</tr>
<tr>
<td>S X D</td>
<td>2/234</td>
<td>.0464</td>
</tr>
<tr>
<td>S X C</td>
<td>2/234</td>
<td>.3560</td>
</tr>
<tr>
<td>All higher order interactions</td>
<td>16/234</td>
<td>.4644</td>
</tr>
</tbody>
</table>

a All three and four factor interactions were tested as a group and were still not significant

* p<.01
Chapter 5

Summary and Conclusions

The effects of stimulus dimensionality and age were the major concerns of this investigation. Hypothesis 1a relating the effect of dimensionality on classification performance was not supported by the results on either Level III or Level IV. There are several possibilities for explaining this unexpected result. One study cited previously (Sigel & Shapir, 1966) supports the result that stimulus dimensionality will not significantly affect the performance of middle class subjects given a classification task. Sigel & Shapiro (1966) reported middle class negro subjects could perform equally well with actual objects or pictorial representation of those objects. A significant stimulus dimensionality effect was not supported by the results of this investigation and all subjects were from middle class homes. According to the results of Sigel & Shapiro (1966) the non-significant stimulus dimensionality effect could be attributed to the socioeconomic status of the subjects. This conclusion is contrary to much of the literature cited in this paper (Dornbush & Winnich, 1966; Etaugh & Van Sickle, 1971; Everett & Armstrong, 1968; Falk, 1968; Gibson, 1963; Stevenson & McBee, 1958; Swayze, 1967; and Tanaka, 1968). Alternative explanations are therefore suggested.

One alternative explanation is concerned with the suitability of the GEBST as an instrument sensitive enough to measure the differences in classification performance due to differences in stimulus
dimensionality. Before this investigation was started the GEBST was postulated to have an attenuated scoring scale for use in this investigation, and the scale was extended in order to increase discrimination of performance at the top end of the scale. After examination of data from this investigation, it seemed that there still may exist an area where the scoring procedure did not adequately reflect the subjects' performance. This seemed to occur in the actual "procedure of grouping" rather than the finished scored results. For example, some subjects seemed initially to be working at a primitive trial and error or pairing procedure, but eventually would reach the correct solution for a one-concept sort. Other subjects could immediately perceive the relationships of a one-concept sort and quickly arrange the stimuli for a one-concept sort. Both of these performances would receive the same score for classification. Inhelder & Piaget (1964) discussed the developmental sequence from trial and error groupings to perception of the entire relationship instantaneously. It is suggested by this researcher that subjects may be working at different classification stages with different dimensions, and this may not be reflected in the performance scores. Though this investigation did not indicate that dimensionality affects classification performance, a new scoring scale or a different classification task might produce different results.

A third possibility to explain the non-significance of dimensionality could be attributed to the inherent design of the GEBST. At Level IV, subjects chose between two characteristics for sorting; an example would be color versus size. The task is thus a choice paradigm and response factors determined by the paradigm could obscure response factors due to dimensionality. The effects
of dimensionality on performance scores could be obscured if, for example, a subject was unable to perform size categorization with pictorial representation. When the choice between color and size was presented, the subject classified pictorial stimuli by color. The classification performance score on this trial would not reflect the difficulty this subject encountered in sorting pictorial stimuli by size. The GEBST does not give a "clear" measure of the variables presented in this investigation, particularly "characteristics." There could also be some undetected interactions between dimensionality and characteristics such as the example just described. Level III should present a "clear" measure of all the variables studied in this investigation, however the task was so easy that variability was obscured due to the "ceiling effect." To assess whether design factors inherent in the GEBST obscured response factors due to dimensionality, it would be necessary to partially replicate this investigation employing alternative measurement procedures to determine whether results occurred as a function of method variance.

It was also found that most scores were very closely grouped. Because there was little variance in the scores, an exceptionally poor or good classification performance had the effect of drastically changing the performance scores of a group in the direction of the unusual score. This restricted distribution of scores could have occurred because the scoring scale did not adequately reflect the classification performance that existed in the sample, and thus extreme scores affected the total performance out of proportion to the number of such scores. With the large resultant variability
within groups due to extreme scores, variability between groups would appear less. A significant between age group variability could thus be masked by a few extreme scores. The number of subjects in a cell was very small and this could further increase the distortion due to one extreme score.

A final possible explanation of the non-significance of dimensionality regards the effects of sampling subjects: their number, age range, and method of selection. The number of subjects in a cell was small (only six subjects in each age group responded to each level of stimulus dimensionality), and a large "n" would have been preferable. The reason this could not be done was the testing time involved with more subjects, and also the problems that occurred in obtaining subjects ages 3 and 4. The gradual increase in performance score with increasing age seems to indicate that the age range selection was adequate. Method of subject selection with all the 3 year olds and most of the 4 year olds was determined primarily by the characteristics of the cooperating schools as almost the entire population participated in this investigation. At the elementary school heterogeneous class groupings existed, and class teachers stated that the subjects presented a typical range of abilities. All ages came from middle socioeconomic status groups so it was thought that subject bias was similar over all ages. A least-squares regression analysis corrected for the arbitrary age grouping, and eliminated the problem of "levelling" on the age variable which occurred in the analysis of variance. Similar results were obtained in both analyses with respect to age.

Though many possible reasons were offered for the non-signifi-
cance of dimensionality, this researcher felt that there were two major explanations. The scoring of the GEBST, as used in this investigation, did not appear to completely reflect classification performance. The second explanation concerns the inherent design of the GEBST. The choice paradigm on Level IV could obscure some of the variability. Though the non-significant result was attributed to the two previous explanations the effect of subject socioeconomic status could not be disregarded. Further research would be required to determine if this was an important factor affecting the results of this investigation.

Hypothesis 2a regarding the effect of subject age on classification was statistically supported by the results of this investigation on both Level III and Level IV. Age was the most significant factor affecting classification performance, and similar results were reported (Graham, Ernhart, Craft & Berman, 1963) on the GEBST. Subject age affected classification performance in the following manner: the youngest subjects received the lowest scores and with increasing subject age there was an increase in performance scores. Hypothesis 2a was supported on Level III and Level IV. The Level III task was too easy for the older subjects, as those of ages 5, 6 and 7 received perfect scores. The use of Level III with older subjects did not contribute as much information about performance on a classification task as was anticipated. The researcher noted that, because of the simple nature of the Level III task, the subjects were put at ease. Subjects began Level IV confident that they could do the task. Younger subjects, who had great difficulty with Level IV were able to feel positive about the test, as they felt that they had at least been able to do the first trials well. Though Level IV contributed almost all the information
concerning classification performance, this researcher felt that Level III was important in creating a confident, positively motivated attitude. For this reason Level III should be retained even with the older subjects. It required no more than five minutes to administer. No subject became bored in this length of time, and not one subject rejected Level III as too simple.

Subjects ages 3 to 7 classified Level IV stimuli with progressive ability. A range of performance existed: the lowest ability level subjects were able to partially perform some trials and the highest ability level subjects were able to complete two-concept sorts on some trials. No subject was able to make two-concept sorts on all the trials in Level IV, and no subject was completely unable to classify the objects in Level IV. Use of GEBST in the manner applied in this investigation seemed appropriate, given the ages of subjects studied.

The development of classification performance, as described (Inhelder & Piaget, 1964), was reflected in the performance of the subjects in this investigation. Some of the youngest subjects utilized trial and error matching of two objects to see what looked alike, next this matching seemed to be done visually and after that a correct one-concept sort occurred almost instantly. A stage seemed to occur when the subject perceived that there were two concepts with which to categorize, however the subject was able to sort on only one concept at a time. The next developmental order after that appeared to be: partial two-concept sorts, then complete two-concept sorts, and finally the interrelations of the two concepts were perceived by a few subjects on certain trials. A similar pattern of behavior was denoted by Inhelder & Piaget (1964) and Piaget (1953, 1954) as preconceptual, and
the age range of the subjects in this investigation was within those authors' suggested range for preconceptual behaviors.

Minor hypotheses related the effects of characteristics and some of their interactions to classification performance. In Level IV minor hypotheses could not clearly be analysed as subjects were able to select one of two characteristics as basis for their classifications. In Level III the perfect performance of ages 5, 6 and 7 eliminated variability in performance. Analysing characteristics with analysis of variance procedure on Level IV demonstrated stimulus characteristics significantly affected classification performance. The effects of the different aspects of stimulus characteristics (color, form and size) presented problems in analysis, however inferences could be deduced from the data to support the minor hypotheses that were made.

Results from Level IV could be taken to support hypothesis 1b that color and form were easier stimulus characteristics to base classes on than size. When size was not one of the possible choices, the highest mean scores occurred. The lowest mean scores occurred with the form-size choice. It may be inferred from this trend that size was the most difficult stimulus characteristic on which to base classes. Color appeared less difficult than form, however the interaction of age with characteristic preference made this inference more tenuous. A comparison of mean characteristic performance scores of 3 year old subjects demonstrated that when color was not one of the choices, classification performance was significantly different. A tabulation of the preferred characteristics (Figure 6) clearly indicated 3 year old subjects chose to sort by color in preference to form or size. Older subjects did not demonstrate this color prefer-
ence nor did they demonstrate significantly different performance between choices of form and color. It can be postulated that for the 3 year old subjects, color was a less difficult characteristic than form but this difference did not occur in subjects 4 years of age or older.

Characteristic preference of subjects on Level IV clearly indicated size was the least preferred dimension over all ages when combined with form. An ambiguous result did occur in the size-color sort choice. Younger subjects definitely preferred color, however older subjects selected size in preference to color. Explanations that could be made for this ambiguous result are pure speculation. The following suggestion is given. The task was so simple for older subjects that they made the task more difficult than it actually was by selecting size as a first basis for categorizing. Older subjects perceived the ordering task presented as a stimulus size categorization and overlooked the easier characteristic of color. Further research would be required to support such a tentative explanation of the size preference of older subjects when presented with a stimulus color-size choice. In Level III a significant interaction occurred between characteristics and dimensionality. The youngest subjects had great difficulty sorting pictures by size. This supports the hypothesis that size is a more difficult characteristic to base sorts on than color or form.

Minor hypothesis 2b stated that size based classifications would be most difficult with pictorial representation and less difficult with three and two dimensional objects. This hypothesis was supported on Level III. On Level IV a significant dimensionality by
characteristic interaction did not occur. The explanation for this could rest with the type of task presented in Level III and Level IV. In Level III the subject was forced to sort pictorial representation by size. In Level IV a subject had a choice of sorting by size or another characteristic. The difficulty of size-based sorts with pictorial representation could be eliminated by the subjects' choice in Level IV (this difficulty was discussed with the problems occurring due to the inherent design of the GEBST). The task presented on Level IV was entirely different from that on Level III. The only test of minor hypothesis 2b occurred with Level III, and it was statistically supported that size-based classifications are most difficult with pictorial representation.

Minor hypothesis 3b was concerned with the interaction of subject age and stimulus characteristics. A significant interaction effect occurred between these two variables with analysis of variance procedures. Hypothesis 3b stated the direction of the interaction: size-based classifications would be more difficult than color or form for the youngest subjects, but with an increase in subject age, there would be smaller differences in performance scores between the characteristic-based sorts. The youngest subjects on Level IV (age 3 years) exhibited exceptionally poor classification performance on the trial with a size versus form sort. Ages 4 to 7 were able to perform in a similar manner on all trial choices. Hypothesis 3b was supported by the results on Level IV. On Level III the lack of variability in performance scores could explain the non-significance of the age by characteristic interaction.

Related research clearly indicated that form was the preferred
dimension when compared with color in all but the youngest subjects (Corah, 1966; Harris, Schaller & Mitter, 1970; Kagan & Lemkin, 1961; and Modreski & Goss, 1969). The results of this investigation supported these findings. Conflict occurred in the results of different researchers over the preference of color in the youngest subjects. The data of this investigation indicated that the youngest subjects preferred color in preference to form as a basis of categorization. Similar findings were reported (Corah, 1964; Suchman & Trabasso, 1966; and Trabasso, Stave & Eichberg, 1969). Size was the least preferred characteristic to use in the classification tasks presented in this investigation. This was supported by the results of Denney (1972b), Kagan & Lemkin (1961); and Sigel (1964). The results of this investigation support the findings of other research. To adequately describe the effects of stimulus characteristics an alternative task to Level IV would have to be developed where subjects did not have a choice between two stimulus characteristics. It was this choice paradigm, implicit in the design of the GEBST that made response factors due to characteristics obscure.

The results of the regression analysis supported the generally accepted finding of Denney (1972a & 1972b), Everett & Armstrong (1968), and Harris, Schaller & Mitter (1970) that in free classification tasks there are no significant sex differences in performance.

In summation, age is the most significant factor affecting performance on a free classification task. The age range of this investigation--3 to 7 years inclusive--seemed adequate to test the performance of preconceptual classification behavior as described by Inhelder & Piaget (1964). The sex of subjects did not significantly affect
classification performance, a finding supported by most research in this area. The characteristics of the stimulus object—color, form and size—was a significant factor affecting classification performance. Size appeared to be a much more difficult characteristic to base classification on than color or form, particularly with the youngest subjects.

Stimulus dimensionality was not found to be a significant factor affecting classification performance. This was contrary to research in this area and also the hypotheses made in this investigation. Two explanations seemed plausible for the results occurring in this investigation: the scoring system of the GEBST did not sufficiently reflect the performance of the tested subjects, and the choice paradigm of Level IV, implicit in the design of the GEBST, obscured some of the variability of stimulus dimensionality.

Further research in the area of stimulus dimensionality in a free classification task seems necessary as the results of this investigation did not support most research which varied stimulus dimensionality. Sigel & Shapiro (1966) reported no stimulus dimensionality effect occurred with middle class subjects. This could explain the results of this investigation. The question still arises that the results of Sigel & Shapiro (1966) are contradictory to all other cited research. Further research into the effect of subject's socio-economic status in the area of stimulus dimensionality seems warranted. Dimensionality has not been fully studied and it has great practical applicability to the study of young children's performance on conceptual tasks. What is needed before this investigation could adequately test the effects of stimulus dimensionality and characteristics is the dev-
elopment of a new free classification test or modification of two areas of the GEBST. To use the GEBST a new scoring scale would need to be developed which would assess the procedures used in groupings as well as the finished classified arrangement. A task needs to be developed similar in difficulty to Level IV which would eliminate the choice paradigm that occurred on Level IV. A subject would then be forced to sort on each dimension with each characteristic. This type of task did occur on Level III but these trials were too simple for most of the subjects in this investigation.

The GEBST was very easy to administer, was enjoyed as a game by the young subjects, took a very short time to administer and was suitable for children ages 3 to 7 years. It was felt by this researcher that free classification tests such as the GEBST could successfully be used to assess conceptual ability in preconceptual children. Research is required to develop a more suitable test. Other investigators have alluded to the advantages of a free classification task to assess conceptual ability in young children but a standardized test for this purpose has not yet been developed.
References


Appendix A

Graham-Ernhart Block Sort Test Manual
4-YEAR PSYCHOLOGICAL EXAMINATION
MANUAL FOR THE
GRAHAM-ERNHART BLOCK SORT TEST
(For Form PS-21)

THE COLLABORATIVE STUDY OF CEREBRAL PALSY, MENTAL RETARDATION AND
OTHER NEUROLOGICAL AND SENSORY DISORDERS OF
INFANCY AND CHILDHOOD

August 1963
4-YEAR PSYCHOLOGICAL EXAMINATION

MANUAL FOR THE GRAHAM-ERNHART BLOCK SORT TEST
(For Form PS-21)

Introduction

The Graham-Ernhart Block Sort test has been included in the 4-year COLR battery as a supplement to the Stanford-Binet (Form L-M). It will increase our sampling of concept formation through the use of a sort where materials vary in color, size and shape. Dr. Graham's previous work with this test indicates that it not only provides a significant discrimination between brain damaged and non-brain damaged preschoolers, but also that it adds to the discrimination between these groups to a degree beyond that possible with the Stanford-Binet alone. (See Table of Reference).

The instructions presented here are a revision of Graham's procedure. The original manual has been altered when it was deemed that such revision would better serve the purposes of the COLR Project without distorting the essential characteristics of the test. We wish to thank Drs. Graham and Ernhart for their generous permission to permit COLR use of the test, and Dr. Graham for her review of this revision.

Materials

The materials consist of 26 10mm. thick plastic blocks in various combinations of three colors (white, red and blue), three forms (circle, square and equilateral triangle), and three sizes (small, medium and large). Areas of the three forms are approximately equal for each size level. The size levels are in the ratio of 1:2:3, the relevant dimension being the sides of the square and the triangle and the diameter of the circle. The exact dimensions are 22, 44 and 66 mm. for sides of the three squares; 33, 67 and 100 mm. for the sides of the three triangles; and 25, 50 and 75 mm. for the diameters of the three circles.

<table>
<thead>
<tr>
<th>Circles</th>
<th>Squares</th>
<th>Triangles</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 large white</td>
<td>2 large white</td>
<td>2 large white</td>
</tr>
<tr>
<td>2 large red</td>
<td>1 large red</td>
<td>1 large red</td>
</tr>
<tr>
<td>2 large blue</td>
<td>1 large blue</td>
<td>1 large blue</td>
</tr>
<tr>
<td>2 medium white</td>
<td>1 medium white</td>
<td>1 medium white</td>
</tr>
<tr>
<td>1 medium red</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 medium blue</td>
<td>1 small white</td>
<td>1 small white</td>
</tr>
<tr>
<td>2 small white</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 small red</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 small blue</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Materials for the Graham-Ernhart test and necessary replacement parts will be furnished by the Central Office.

General Administration Instructions

1. This test is given as a whole. Other procedures should not be interposed between trials or levels. Such interposition, if practiced, might produce unknown interference with establishing the proper sorting set.

2. The Graham-Ernhart Block Sort test items are organized into four levels. Level III is always given first. Depending upon the child's performance on Level III, either Level II or Level IV, but not both, is given. Graham's Level I is not being used by the COLR Project, since it is anticipated that Level I is not needed to provide satisfactory discrimination between subjects at the lower ranges.

3. Throughout the test, the examiner should avoid the use of phrases relevant to concepts of size, shape, or color in giving his instructions; e.g., "little ones," "blue blocks," "circles." Stick to use of phrases and words such as "look alike," "alike," "same," and "belong together."

4. Since the same concepts are tested repeatedly, the examiner should not indicate that a correct response has or has not been made, except on the three trials of Level III. However, if the child, having made a correct sort, then proceeds to build or play at random with the blocks (thus destroying the
sort) the examiner should, if possible, say before the sort is destroyed, "Is that the way you want them?", or, "Have you finished putting the blocks that are the same together?", and proceed to score the sort in terms of the block placement at that point.

5. If a child requests help, the examiner should repeat the instructions or say, "Just do the best you can."

6. Blocks not in use should be kept out of sight.

7. In scoring, verbal recognition is not counted. Only the child's nonverbal sorting behavior contributes to his score.

8. To avoid supplying extraneous cues, the examiner should hold all the required blocks in his hand before placing any of them on the table.

9. At all levels, each trial is presented only once.

Administration and Scoring:

Definitions:

A. A group is two or more blocks which are adjacent to one another (they may or may not actually touch), and are spatially separated from other blocks. All blocks included in a trial may form a single group. The definition of group demands only that two or more blocks are together in a stack, row, or laying together. On occasion a group of blocks may contain combinations of stacks or rows, i.e., be both horizontally and vertically arranged.

B. A subgroup is two or more blocks within a group which are alike in size, form or color, whichever dimension(s) is (are) being varied on a given trial.

NOTE: The distinction in definition between a group and subgroup is critical to the scoring system and needs to be thoroughly understood. Blocks in a group may or may not be alike; the only essential requirement is that they be clustered. The definition of a subgroup implies that the group within which these like blocks are found includes other subgroups or misplacements. The blocks within a specific subgroup are always alike by definition. Blocks constituting a subgroup must also be clustered.

In Example I, the cluster of three blocks constitutes a group; the cluster of two blocks is also a group. Within the three block group is a subgroup of two circles which are alike in size. The two white blocks in the cluster of four form a subgroup. The cluster of four blocks forms an additional group.

Example I:

```
B: Blue
R: Red
W: White
```

In Example II the cluster of three blocks is a group; so are the two clusters of two blocks each.

Example II:
In Example III, the three blue circles within the group of five constitutes a subgroup of three blocks. Other blocks in the diagram are extraneous to this example.

Example III:

Finally, note that a particular block may belong to more than one subgroup. In Example IV, the middle blue square forms a subgroup of squares with the block on its left and also forms a subgroup of two blue blocks with the block on its right. In the first instance the subgroup is formed on the basis of shape, in the second on the basis of color.

Example IV:

C. A leftover block is one not adjacent to any other block. The blocks called extraneous in Example III are leftover blocks.

D. Misplacements are blocks within a group so placed that they form no subgroup. They are not similar to blocks next to them in the group, in terms of the concepts being sorted (shape, color, or size). In Example III, the white block and the red block in the 5 block group are misplacements.

Level III: Administration and Scoring

1. The task is to sort six blocks into three groups of two identical blocks. The blocks used and the orders in which they are placed on the table before the child are given below. The order of placement is specified from the examiner's point of view as he sits opposite the child.

Trial 1: Six large circles

Red White Blue Red White Blue

Child

Examiner

Trial 2: Six white circles

Small Medium Large Small Medium Large

Child

Examiner

Trial 3: Six large white blocks

Circle Square Triangle Circle Square Triangle

Child

Examiner
2. After the blocks have been placed in a row on the table, say "Put the ones that (are alike, are the same, or belong) together." This instruction can be repeated as necessary.

3. A completely correct sort on a given trial consists of pairing the identical blocks with each other. Such a sort earns two points.

4. A partially correct sort on a given trial is one where the sort is incomplete but where the concept is apparent both in the sorting behavior and in the final arrangement. A sort with only one or two misplacements or leftovers or a sort in which two identical blocks are placed together but apart from the remaining blocks are examples of partial sorts. Note that the definition of a partially correct sort includes the specification that the examiner must see some evidence of concept control in the sorting behavior prior to the final arrangement. This clinical criterion has been added to avoid the crediting of final arrangements where two identical blocks happen by chance to end up closer to each other than they are to other blocks. A partially correct sort earns one point.

5. Examples of Level III, Trial 3 sorts.

For the purposes of scoring, stacking or placing blocks in a row will be considered equally correct.

A. Correct sort; earns two points.

B. Correct sort; earns two points.

C. Partially correct sort; earns one point. Partly correct because one group pairs identical mates, even though there are four misgrouped blocks.

D. Partially correct sort; earns one point. There are only two misplacements (the two triangles).

E. Partially correct sort; earns one point. There are only two misplacements (the two triangles).
Partly correct sort. There is one misplacement and one leftover (the two triangles).

Incorrect sort; zero points. Three blocks are misplaced and one is leftover.

Incorrect sort; zero points. There is a joined identical pair, but they are not spatially discrete from the other four misgrouped blocks.

6. To encourage more discrete grouping on subsequent trials at this and other levels, additional instructions and demonstration are given following each of the trials at this level:

A. If the child achieves a two-point sort and has arranged the blocks into three discrete groups say, "Fine," and go on to the next trial.

B. If the child achieves a two-point sort but has not arranged the blocks into three discrete groups (for example, stacking all six blocks into a single pile, or a single row) say, "Fine, these two are alike, and these two are alike, and these two are alike" as you separate the groups.

C. If the child achieves partial or no credit say, "Let me show you. These two are alike, and these two are alike, and these two are alike," as you rearrange the blocks into three separate groups. Do not allow the child another try at the same trial after it has been demonstrated.

7. Maximum score on Level III is six points.

Administration Choice-Point

1. If the child's score on Level III is two points or less, administer Level II. In this case Level IV is not administered and no credit for it is given.

2. If the child's score on Level III is more than two points, administer Level IV and give the child an automatic credit of 9 points for Level II.

Level II: Administration and Scoring

1. The child's task at this level is to choose the mate to the block held by the examiner from the three blocks placed before him. The block the child is to match is held about one foot above the middle of the three blocks on the table, and the examiner says, "Point to the one that looks just like this one."
Level II: Administration and Scoring (Cont.)

2. The order of the blocks placed on the table for each trial are specified from the examiner’s point of view as he sits opposite the child facing him.

   Trial 1: On the table: 3 large circles; Red  White  Blue
           Examiner’s order of presentation: Blue  Red  White

   Trial 2: On the table: 3 white circles: Small  Medium  Large
           Examiner’s order of presentation: Small  Large  Medium

   Trial 3: On the table: 3 large white blocks: Circle  Square  Triangle
           Examiner’s order of presentation: Square  Triangle  Circle

3. Every effort should be made to ensure having the child's attention before each request for him to point. Once given, an incorrect point should be accepted at face value.

4. Each correct match earns one point, making a maximum score of three per trial and a total maximum of nine points for Level II.

5. If child is credited with the full 9 points on Level II, administer Level IV as part of the test. However, the points earned on Level IV should not be included in the score for summary sheet purposes, but will be analyzed at a future date.

Level IV: Administration and Scoring

1. The task for each trial of Level IV is to sort the nine blocks, which on any given trial vary along two of three possible dimensions (size, color, and form) into groups of like blocks.

   General Instructions

   The instructions are the same for each of the three trials. The blocks used in a given trial are spread out on the table in random order. The blocks should be spread over a roughly oval area, not in a row. The examiner says, “Now put the ones that (are alike, are the same, or belong) together.”

   Blocks to be used for specific trials are as follows:

   Trial 1: Color and size vary with form held constant. Use nine circular blocks with one of each size and color.

   Trial 2: Color and form vary with size held constant. Use nine large blocks with one of each form and color.

   Trial 3: Size and form vary with color held constant. Use nine white blocks with one of each size and form.

   Scoring of Level IV takes only final arrangement into account. By following the rules below, the number of errors on a given trial can be determined. Since it is unlikely that the examiner will have these rules well enough in mind, especially in his early use of the test, it is mandatory that the examiner sketch the final arrangement on each trial, and then score it later.

   Scoring Rules

   A. Any group of blocks which are all alike on one of the varied concepts is considered correct.

   B. Any subgroup of three blocks is always correct.

   C. Subgroups of two blocks are counted as errors (one point per block) if:
**Level IV: Administration and Scoring. (Cont.)**

1. the group contains a misplacement, or
2. the group which contains the subgroup of two blocks is composed of only three blocks.

**Note:** Groups of three blocks which are **not** all alike and which contain two subgroups of **two** blocks per subgroup are errors by rule C2. They constitute three error points, one for each block. (see Example C, p. 8).

D. **Subgroups** of only two blocks are correct if all other blocks in the group also form subgroups, **and** the group consists of more than three blocks.

E. **Leftovers** and misplacements count one error per block.

F. Typically, sorting by four-year-olds is on the basis of one varied concept or the other, **or** some use of both concepts. Thus, one group may be formed having blocks of the same color, **and** another group in the same trial will contain blocks of the same form or size. It is possible but rare for a child to produce a perfect sort on both concepts. To do so, **every** block must participate simultaneously in two subgroupings. This can be achieved with a $3 \times 3$ matrix in which blocks of one concept form the rows, and blocks representing the other concept form the columns. It can also be achieved with three stacks formed on the basis of one concept with the second concept demonstrated by uniform positions within stacks in all three stacks.

G. Errors on a given trial may range from 0 to 9. They should be converted to points toward the total raw score by use of the table below. Maximum raw score points possible on a given trial is ten. Maximum raw score contributions of Level IV to total raw score is 30.

<table>
<thead>
<tr>
<th>Errors on a trial</th>
<th>Raw Score for the trial</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>0</td>
<td>9</td>
</tr>
<tr>
<td>1</td>
<td>8</td>
</tr>
<tr>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td>0</td>
</tr>
</tbody>
</table>

**Scoring Examples**

This section of scoring examples should be studied carefully prior to scoring of actual protocols. You may find it helpful to color the sketches. Additional examples of difficult scoring problems with commentary will be added as our project experience with this test accumulates.

In the examples below, the letters R, W, and B are used for the three colors — red, white, and blue. Large, medium and small size is indicated by the relative size of the drawings. All examples are drawn as if arranged on a horizontal plane, **but** these arrangements may be produced by stacks **or** by a three dimensional pattern involving both horizontal and vertical groups.

Example A. Color and Size Problem. **Nine errors.**

The subgroup of two within the three group is not considered correct (rules C1 and C2). Other errors are four leftovers and a paired misplacement (rule E).

---

August 1963
Example B. Color and Form Problem. Nine errors.
The two two-member subgroups of like form are part of a group containing misplacements (rule C₁).

Example C. Color and Form Problem. Nine errors.
Two two-block subroups in a group of only three blocks are not considered correct by rule C₂.

Example D. Color and Size Problem. Six errors.
The group of three is considered correct by rule A.
The remaining blocks are leftovers (rule E) or paired misplacements (rule E).

Example E. Color and Size Problem. Six errors.
Four leftovers and two misplacements. The subgroup of three blue circles is correct by rule B.

Example F. Color and Form Problem. Five errors.
The two groups of two are correct by rule A. The errors are the group of five misplacements.
Example G: Color and Size Problem. Five errors. Five leftovers. Note that the group of four blocks has been arranged into two subgroups each containing two blocks of like color and one subgroup of two blocks of like size. The two middle blocks have been sorted simultaneously on both concepts. Group of four correct by rule D.

Example H: Color and Size Problem. Five errors. Three leftovers and a pair of misplacements. The middle group of four blocks is correct by rules B and D. This group involves one subgroup sorted by color, another subgroup sorted by size, and one block sorted simultaneously on both bases.

Example I: Form and Size Problem. Four errors. One leftover, one misplacement, and one paired misplacement. The correct group of two qualifies by rule A; the correct subgroup of three by rule B.

Example J: Color and Size Problem. Four errors. Two leftovers, two misplacements, Rules A and B make the other five blocks correct.
Example K. Color and Size Problem. Four errors.
The two subgroups of three on the left side of the row (with one block sorted simultaneously on both concepts) are correct by rule B. The two-member subgroup of blue circles is incorrect by rule C and the other two blocks are misplacements (rule E).

Example L. Size and Color Problem. Three errors.
All leftovers. The groups of three are correct by rule A.

Example M: Color and Form Problem. Three errors.
The errors are represented by a group of three which contains a subgroup of two and a misplacement (rule C). The row-group of six blocks contains two subgroups of three, and one subgroup of two, with two blocks in the row sorted simultaneously on two concepts and hence having representation in two subgroupings. The entire row-group is correct by rules B and D.

Example N. Color and Size Problem. Two errors.
A misplacement and a leftover. The two-member groups are correct by rule A; The three-member subgroup by rule B.
Example O. Color and Size Problem. One error. (Rule E). The others correct by rule A.

Example P. Form and Size Problem. Zero errors. The two-member square group and the three-member triangle group are correct by rule A. The row of four blocks is correct by rules B and D. It includes a subgroup of three circles and another of two blocks of like size.

Example Q. Form and Size Problem. Zero errors. Three-block subgroups are always correct (rule B) and two-block subgroups are correct in this case because they fit rule D.

Example R. Color and Size Problem. Zero errors. An example of a perfect sort involving sorting simultaneously on both concepts. There are six subgroups of three blocks, all correct by rules B and F.
Table of References


Appendix B

Consent Form Letter
Appendix C

Pictorial Representational Stimuli
Pictorial representational stimuli—scale 2/3 actual size