INFORMATION TO USERS

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

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

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

2. When an image on the film is obliterated with a round black mark it is an indication that the film inspector noticed either blurred copy because of movement during exposure, or duplicate copy. Unless we meant to delete copyrighted materials that should not have been filmed, you will find a good image of the page in the adjacent frame. If copyrighted materials were deleted you will find a target note listing the pages in the adjacent frame.

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

4. For any illustrations that cannot be reproduced satisfactorily by xerography, photographic prints can be purchased at additional cost and tipped into your xerographic copy. Requests can be made to our Dissertations Customer Services Department.

5. Some pages in any document may have indistinct print. In all cases we have filmed the best available copy.

University
Microfilms
International

300 N. ZEEB RD., ANN ARBOR, MI 48106
CHILDREN'S ABILITY TO CONCEPTUALIZE CONFLICTING EMOTIONS:  
A COGNITIVE-DEVELOPMENTAL STUDY  
DISSERTATION  

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

By  
Bruce Hirsch, B.A., M.A.  

* * * * *  
The Ohio State University  
1981  

Reading Committee:  
Charles Wenar, Ph.D.  
Fred Damarin, Ph.D.  
John Gibbs, Ph.D.  

Approved By  
Department of Psychology
ACKNOWLEDGMENTS

This research was facilitated by a number of people, and their help is gratefully acknowledged.

My adviser, Dr. Charles Wenar, and the other members of my committee, Dr. Fred Damarin and Dr. John Gibbs, have offered helpful suggestions at various stages of this work. Dr. Susan Harter of the University of Denver, whose research was the impetus behind this work, provided encouragement and information and helped me to clarify my ideas in the crucial beginning period of this study.

The cooperation of many people at Sam-Kel Children's Academy (pilot study) and Bexley City Schools (main study) greatly facilitated my work. At the Children's Academy, I would especially like to thank Barbara Sams, Anita Mason, and Lori Marquard.

Dr. Loren Thompson provided the initial approval for my work in the Bexley Schools. At Montrose Elementary School, the principal, Ms. Linda Lucks, was very helpful in coordinating my work at her school. I would also like to thank all of the teachers at Montrose for their flexibility, encouragement, and especially for their hospitality. Mr. Jack Brandt (principal) and Ms. Nancy
McCreary (counselor) helped my brief work at Cassingham Junior High School to run very smoothly.

A special thank you goes to the parents for granting permission and to the children for their willing cooperation and pleasant manner.

Finally, I thank my wife, Paula, for her willingness to make sacrifices so that this dissertation could be completed and, most of all, for her reliability.
VITA

August 24, 1949.... Born—Brooklyn, New York
1971............... B.A., Yale University, New Haven, Connecticut
1974-1976......... Teaching Associate, Department of Psychology, The Ohio State University, Columbus, Ohio
1977.............. M.A., The Ohio State University, Columbus, Ohio
1977-1978........ Psychology Intern, Pittsburgh Child Guidance Center, Pittsburgh, Pennsylvania
1978-1980........ Psychology Assistant, Children's Mental Health Center, Columbus, Ohio
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>ii</td>
</tr>
<tr>
<td>VITA</td>
<td>iv</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>vi</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>viii</td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>REVIEW OF THE LITERATURE</td>
<td>7</td>
</tr>
<tr>
<td>The Cognitive-Developmental Model</td>
<td>7</td>
</tr>
<tr>
<td>Invariant Sequence</td>
<td>11</td>
</tr>
<tr>
<td>Overall Structure</td>
<td>23</td>
</tr>
<tr>
<td>The Social and Affective Domain</td>
<td>44</td>
</tr>
<tr>
<td>The Ability to Conceptualize Conflicting Emotions</td>
<td>55</td>
</tr>
<tr>
<td>METHODS</td>
<td>66</td>
</tr>
<tr>
<td>Subjects</td>
<td>66</td>
</tr>
<tr>
<td>ACCE Interview</td>
<td>67</td>
</tr>
<tr>
<td>Measures of Logico-physical Reasoning</td>
<td>75</td>
</tr>
<tr>
<td>Overall Procedure</td>
<td>81</td>
</tr>
<tr>
<td>RESULTS</td>
<td>84</td>
</tr>
<tr>
<td>Overview</td>
<td>84</td>
</tr>
<tr>
<td>Scalogram Analysis</td>
<td>91</td>
</tr>
<tr>
<td>ACCE as an Overall Structure</td>
<td>96</td>
</tr>
<tr>
<td>ACCE and Piagetian Measures of Concrete Operations</td>
<td>108</td>
</tr>
<tr>
<td>DISCUSSION</td>
<td>114</td>
</tr>
<tr>
<td>SUMMARY</td>
<td>137</td>
</tr>
<tr>
<td>APPENDIX</td>
<td>142</td>
</tr>
<tr>
<td>Rules</td>
<td>142</td>
</tr>
<tr>
<td>Examples</td>
<td>145</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>152</td>
</tr>
</tbody>
</table>
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Percentage of Subjects Passing Each Category of Response for Overall Scores..</td>
<td>85</td>
</tr>
<tr>
<td>2. Percentage of Subjects Passing Each Category of Response for Happy.</td>
<td>85</td>
</tr>
<tr>
<td>3. Percentage of Subjects Passing Each Category of Response for Sad.</td>
<td>85</td>
</tr>
<tr>
<td>4. Percentage of Subjects Passing Each Category of Response for Mad.</td>
<td>86</td>
</tr>
<tr>
<td>5. Percentage of Subjects Passing Each Category of Response for Scared.</td>
<td>86</td>
</tr>
<tr>
<td>6. Analysis of Variance Summary.</td>
<td>89</td>
</tr>
<tr>
<td>7. Scalogram Analysis by Emotions.</td>
<td>92</td>
</tr>
<tr>
<td>8. Scalogram Analysis for Overall Scores.</td>
<td>92</td>
</tr>
<tr>
<td>9. Interitem Homogeneity Coefficients for Pairs of Response Categories.</td>
<td>95</td>
</tr>
<tr>
<td>10. Independence of Responses for Each Emotion Analyzed Using Frequencies of All-or-none &amp; Mixed Scores.</td>
<td>97</td>
</tr>
<tr>
<td>11. Bivariate Frequencies for Happy &amp; Sad (All Subjects).</td>
<td>99</td>
</tr>
<tr>
<td>12. Bivariate Frequencies for Happy &amp; Mad (All Subjects).</td>
<td>99</td>
</tr>
<tr>
<td>14. Bivariate Frequencies for Sad &amp; Mad (All Subjects).</td>
<td>100</td>
</tr>
</tbody>
</table>
15. Bivariate Frequencies for Sad & Scared (All Subjects) ........................................... 100
16. Bivariate Frequencies for Mad & Scared (All Subjects) ........................................... 100
17. Measures of Association for Levels of Pairs of Emotions (All Subjects) ..................... 101
18. Bivariate Frequencies for Happy & Sad (Ages 5-6 & 7-8) ...................................... 102
19. Bivariate Frequencies for Happy & Mad (Ages 5-6 & 7-8) ...................................... 102
20. Bivariate Frequencies for Happy & Scared (Ages 5-6 & 7-8) .................................... 102
21. Bivariate Frequencies for Sad & Mad (Ages 5-6 & 7-8) ........................................ 103
22. Bivariate Frequencies for Sad & Scared (Ages 5-6 & 7-8) ...................................... 103
23. Bivariate Frequencies for Mad & Scared (Ages 5-6 & 7-8) ...................................... 103
24. Measures of Association for Levels of Pairs of Emotions (Ages 5-6 & 7-8) ............... 104
25. Percent of Subjects at Each Overall Level of ACCE Showing Different Maximum Discrepancies Between Levels for Any Two Emotions (All Subjects) ......................... 107
26. Percent of Subjects at Each Overall Level of ACCE Showing Different Maximum Discrepancies Between Levels for Any Two Emotions (Ages 5-6 & 7-8) .......................... 107
27. Correlations Between ACCE and Piagetian Tasks ...................................................... 109
28. Bivariate Frequencies for ACCE & Conservation ....................................................... 110
29. Bivariate Frequencies for ACCE & Each Classification Task ..................................... 112
30. Measures of Association Between ACCE & Classification Tasks ............................... 112
**LIST OF FIGURES**

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Scores for each category summed across emotions and subjects for each age group</td>
<td>90</td>
</tr>
</tbody>
</table>
INTRODUCTION

It has become increasingly popular to apply a cognitive-developmental approach to a wider and wider range of content areas. Particularly in recent years, this approach has been found helpful in studying children's understanding of social and affective issues, as well as what has been thought of traditionally as the cognitive domain. Harter's (1977) paper on children's conceptualization of conflicting emotions and her more comprehensive study of children's understanding of multiple emotions (Harter, 1979) are part of this trend.

As the following Review of the Literature will discuss in greater detail, the two principal themes of the cognitive-developmental approach are sequence and structure. More specifically, the theory predicts that each individual will attain successive levels or stages of development in the same order as every other individual, although rates of development may vary. The notion of structure refers to the formal characteristics of a task as opposed to the specific content of the task and implies that tasks with the same underlying structure should be attained at the same time. To anticipate a conclusion of that Review, we
can state that although these hypotheses have not always been empirically verified, they remain valuable in stimulating and guiding research. In addition, the structural hypothesis, which has received less empirical support, has been revised by some writers, so that some would now maintain that structurally related tasks do not need to develop in strict synchrony to accord with cognitive-developmental theory. Nevertheless, all cognitive-developmentalists consider the structural aspect of a task important, although some feel that content is also important.

These theoretical concepts can be translated into Harter's work and the present study as follows. Harter (1979) identified three dimensions which can be considered the structural variables in the conceptualization of multiple emotions. The temporal dimension refers to whether the emotions are thought of as sequential or simultaneous. The object dimension refers to whether different emotions are thought of as directed toward the same object or different objects. The valence of an emotion refers to whether it is a positive feeling or a negative feeling, and for multiple emotions the structural dimension is whether both emotions are of the same valence or of opposite valence. Content in the conceptualization of multiple emotions would be the particular emotions dealt with.

In regard to the hypothesis of an invariant sequence, Harter (1979) found support for the idea that the
conceptualization of sequential emotions always precedes the conceptualization of simultaneous emotions. She also presented a possible sequence involving all eight combinations of the temporal, object, and valence dimensions. However, this more differentiated sequence was considered preliminary and tentative because only the temporal dimension had been systematically investigated. The present study will systematically explore the object dimension as well as the temporal dimension. Furthermore, the valence dimension will be controlled in that this study will only look at conflicting emotions, defined as multiple emotions of opposite valence. Thus, this study will attempt to replicate Harter's finding of an invariant sequence regarding the temporal dimension and will also test whether there will still be an invariant sequence when more categories of response are differentiated.

In addition to the question of whether there is an invariant sequence in the ability to conceptualize conflicting emotions (hereinafter abbreviated ACCE), there is the question of just which order this sequence takes. In particular, Harter (1979) and Selman (1976b) both hypothesize an invariant sequence for ACCE but they postulate different sequences. This issue will also be investigated in the present study.

The issue of structure in regard to ACCE has several facets. First of all, there is the question of whether
responses for each emotion will follow the same structurally-defined sequence. In a sense, this is the least that should be expected according to cognitive-developmental theory. Harter's (1979) data indicated that the same three-step sequence of no two emotions conceptualized together, two emotions conceptualized in sequence, two emotions conceptualized as simultaneous was followed for each of the four emotions studied. However, there was some indication that the more differentiated sequences involving the temporal, object, and valence dimensions may be different for different emotions. This issue will receive further scrutiny in the present study.

The second structural question relates to the timing of the achievement of parallel levels for the different emotions. In the strictest sense, cognitive-developmental theory would predict that each individual would reach equivalent levels of reasoning for each emotion simultaneously. More loosely, we would expect a tendency for children who can conceptualize conflict of a particular form involving a particular emotion to be able to do the same for other emotions. If there is not precise synchrony in development in different content areas, an additional question of interest is whether there are periods of consolidation during which the consistency among the levels of ACCE for different emotions increases.
The third facet of the structural issue concerns the relationship between ACCE and Piagetian measures of logico-physical reasoning. In particular, Harter (1979) noted that, like ACCE, conservation and multiple classification skills depend on developing the ability to consider more than one aspect of a situation simultaneously. The synchronous development of tasks which share such structural similarities despite disparity of content is basic to cognitive-developmental theory. We would thus expect at least a significant correlation between ACCE and each of the other measures of cognitive development.

This study will also examine Selman's (1976b) hypothesis that particular levels of logico-mathematical reasoning underly developmental skills in the social and affective spheres. This hypothesis predicts that for equivalent levels, an individual may possess the logico-mathematical skill without the social-affective but not vice versa. Finally, an attempt will be made to determine to what extent we can define equivalent levels of ACCE on the one hand and conservation or classification skills on the other. Overall then, this study is concerned with the question of to what extent and in what ways the development of ACCE fits the cognitive-developmental model.

One point of clarification is necessary. We are not concerned here with the development of emotions, but with
the development of cognitions regarding those emotions. Of course, it seems unlikely that these two attributes would develop independently in the child, but it is possible to separate them logically. This research may be useful as a prelude to integrating the psychology of affective development and the psychology of cognitive development, but such an integration is beyond the scope of this study.
Kohlberg (1969), in discussing the essential features of his theoretical perspective concluded that, "The core of the cognitive-developmental position, then, is the doctrine of stages" (p. 352). Underlying this summary statement was a further explanation of the notion of stages, which had been put forth by the Genevan theorists. At a conference, Inhelder (1956/1971) stated three criteria of stages: "1. The stages of development are defined by structured wholes.... 2. The passage from an inferior stage to a superior stage is equivalent to an integration.... 3. The order of succession of stages is constant..." (pp. 84-85).

At a later session of the same conference, Piaget (1960/1971a, pp. 13-14) elaborated five criteria for stages. In addition to the three already mentioned, he stated that each stage is qualitatively different from the stage before and therefore represents an achievement with respect to the previous stage and a preparation with respect to the following stage. Finally, he defined stages as successive levels of equilibrium.
Frequently, this conceptualization of stages has been seen as involving two essential, separate but interconnecting, aspects of development—sequence and structure (Flavell, 1970; Sigel, 1968; Wohlwill, 1963). Each of these aspects requires some amplification. In cognitive-developmental theory, sequence generally refers to an invariant sequence or a constant order of succession. These stage sequences may involve the overall course of development (such as Piaget's sequence of sensorimotor, preoperational, concrete operational, formal operational), stages within a particular level (such as the six stages of sensorimotor development, Piaget, 1937/1954), or a sequence of steps in a particular domain, such as classification (Inhelder & Piaget, 1959/1964).

In any of these cases, all individuals are seen as passing through the stages in the same order, although their rates of progress may vary. Although some have claimed that this invariance implies a maturationist point of view (e.g. Beilin, 1971), Piaget (1960/1971b, pp. 100-104) has explained these sequences in terms of equilibration, which implies a more interactionist perspective. Although the explanation of these sequences has often been seen as a weak point in the theory, their existence is a crucial element of the model.

The concept of structure has to do with the hypothesis that cognitive gains do not occur in isolated units
but are linked together with other cognitive elements with which they share certain formal characteristics. For Piaget (e.g. 1970), overall structures usually refer to cognitive abilities which possess certain mathematical properties. For example, a structure like classification at the concrete operational level is considered a "grouping" because it has such properties as combination, addition, reversibility, associativity, and an identity element which leaves the class unchanged. Other cognitive-developmentalists, such as Selman (1976b), use the term structure in a looser sense to refer to any organizational characteristics of thought as opposed to the specific content of thought. This definition would include the more mathematical Piagetian conception. In fact, at least according to Flavell (1963, p. 18), Piaget's earlier work reflects this more inclusive definition of structure and he never entirely abandoned this broader notion. In any event, the term structure will be used here in the looser, more inclusive sense.

Examples of structures abound in the Piagetian literature--primary circular reaction, object permanence, one-to-one correspondence, seriation, proportionality, etc. Structures can be relatively broad or relatively specific. For example, we might consider conservation a particular structure. However, we can also view conservation as a specific content area within larger structures, such as the
ability to reverse mental processes and the ability to consider two dimensions of a situation simultaneously, which are two hallmarks of the concrete operational stage. Alternatively, conservation contains within it more specific structures, such as conservation of substance.

In all cases, the concept of structure suggests that certain cognitive attainments, which are linked logically and psychologically, should occur together. This is the complement of the notion of sequence, which suggests that certain cognitive attainments should succeed each other in a fixed order. Cognitive-developmental stages then are overall structures or groups of related structures which occur in an invariant sequence.

As Wohlwill (1963) pointed out, an important characteristic of these hypotheses of sequence and structure is that they are empirically testable. At that time, he lamented that so much research had been devoted to verifying the developmental changes that Piaget had reported and to specifying the environmental conditions that influence those changes, and so little research had been devoted to testing the basic hypotheses of the theory. Since that time, there has been considerable research on both invariant sequence and overall structure as cognitive-developmental hypotheses.
Invariant Sequence

The most straightforward method for documenting an invariant sequence would be a longitudinal study. But perhaps the only longitudinal test of Piagetian sequences was done by Dudek and Dyer (1972). They evaluated subjects yearly over a four-year period on nine Piagetian measures. Each task was scored at a minimum of three stages—preliminary, intermediate, and terminal—and for some tasks the scoring was further subdivided. They found that 10.4% of the responses were regressions from one stage to a theoretically earlier stage. However, they emphasized that only 6.5% of these regressions (or less than 1% of the total responses) were "true" regressions in the sense of being a regression from a terminal stage to a preliminary stage. Thus, they interpreted the results as supporting Piaget's concept of constant, irreversible stages.

It would be interesting to know how many of these regressions actually involved a return to a lower level of thinking after a higher level had been achieved and how many involved a change in the usual sequence, such that the child achieved the supposedly earlier stage for the first time after a supposedly later stage. A longitudinal design could provide such information, but Dudek and Dyer reported no such analysis.

Undoubtedly, the main reason that more longitudinal studies have not been done is that there are so many
practical problems in carrying them out. However, another factor arguing against longitudinal studies is the potential effect of periodic retesting in altering the normal course of development (Wohlwill, 1973). The question of experimental effects is even more significant when the procedure is explicitly designed to foster development. In a sense, such "training" studies can be seen as condensed, short-term longitudinal studies. For example, Smedslund (1961) subjected non-conserving children to conflict situations in order to induce conservation of quantity. Out of 30 children who changed in some way, only 2 did not fit the hypothesized sequence, which was based on pretest results. Of course, the question remains whether normal development would follow the same course as this induced development.

Despite the importance of longitudinal and training studies when they are done, the fact of the matter is that most cognitive-developmental studies have used a cross-sectional, one-time measurement design. It is necessary, before discussing further research on the hypothesis of invariant sequence, to examine the adequacy of various cross-sectional statistical shortcuts to testing this hypothesis.

The most common statistical technique used to test the hypothesis of invariant sequence is scalogram analysis. This technique, originally developed by Guttman (1950) to
test the unidimensional nature of attitude questionnaires, determines to what extent passing more difficult items implies passage of all easier items. Its application to the Piagetian notion of invariant sequence is rather straightforward. If a sequence of cognitive attainments always occurs in the order A, B, C, and if each subject retains the earlier abilities after he or she achieves higher levels, then we would expect all subjects to show one of four scale patterns of passes: no passes, A, AB, ABC. If, however, some subjects show non-scale patterns, like B or AC, this indicates that at least some individuals follow a developmental sequence different from the hypothesized sequence.

A few words of caution regarding scalogram analysis are in order. First of all, in developmental studies, it is appropriate only for what Wohlwill (1973) calls the cumulative case, in which previous abilities are retained. In the disjunctive case, where each level in a sequence replaces the previous level, a scalogram model cannot be used. It is evident that much of Piagetian theory deals with disjunctive sequences, such as no ability to conserve, vacillation between conservation and nonconservation, consistent conservation. Wohlwill pointed out that there are statistical models, equivalent to Guttman scaling, for disjunctive cases, but these are more complex and have generally not been used in developmental research. Longitudinal
designs, like that used by Dudek and Dyer (1972), are particularly appropriate for testing disjunctive sequences. Wohlwill questioned the practice, apparently used in constructing some infant scales, of converting a disjunctive sequence to a cumulative sequence by automatically crediting passes on all lower levels when a higher level is passed. Whatever the merits of such a procedure in scale construction, it is clearly inappropriate as evidence of invariant sequence because it begs the question.

The other caution involving scalogram analysis relates to the interpretation of results. The principal measure of scalability proposed by Guttman (1950) was the coefficient of reproducibility (Rep), which expresses the percentage of item scores that can be reproduced by knowing the scale scores of the subjects. This is based on the notion that if scalability is perfect, we know exactly which items each subject passed if we know how many items that subject passed. Guttman adopted the fairly stringent criterion of 90% scalability (Rep = .90), and although this was somewhat arbitrary, it has been generally accepted. Actually, the problem with this criterion is not its arbitrariness, but the fact that it is not based on probabilities. Depending on the marginal frequencies for each item (i.e. the percentage of subjects passing each level of a skill), the .90 criterion may be very difficult to attain or it may represent only a chance occurrence. Recognizing this,
Guttman recommended paying attention to such factors as the range of marginal distributions and the number of items in assessing whether scalability may have occurred by chance. He also asserted that the pattern of non-scale types should appear random, rather than certain non-scale types appearing with enough frequency to be noticed. This is particularly important for developmental studies, where such clustering of non-scale types may indicate a minority alternative sequence of development.

To deal with the issue of evaluating the reproducibility in terms of the marginal frequencies of the items, Green (1956) proposed an index of consistency (I), which was defined as
\[ I = \frac{(\text{Rep} - \text{Rep}_I)}{(1 - \text{Rep}_I)} \]
where \( \text{Rep}_I \) is the minimum reproducibility that would be obtained, given the marginal frequencies of the items, even if the items were independent of each other. This index varies from 0, when the items are independent, to 1, when there is perfect scalability. Green proposed an I of .50 as an acceptable minimum criterion, and this too has become generally accepted. Like the criterion for Rep, this criterion has no specific probability associated with it.

Statistical probability tests of scalogram analysis have also been developed. Sagi (1959) devised an exact test to determine whether a particular Rep is greater than chance, but his method works only when there are four or
fewer items. Proctor (1970) has developed a chi-square goodness-of-fit statistic to test whether a given set of scale and non-scale types can reasonably be assumed to derive from a scalogram pattern. Since there are a variety of criteria for accepting or rejecting a given sequence as scalable and hence invariant, such a decision is not an automatic process. Rather, the various criteria need to be considered as to their relevance both methodologically and conceptually. In practice, most developmentalists have considered a sequence scalable if the reproducibility is equal to or greater than .90 and the index of consistency is equal to or greater than .50.

A brief comment is also necessary regarding statistical methods other than scalogram analysis that have been used to examine developmental sequences with cross-sectional samples. These have included analysis of variance tests of the effect of age, alone or in interaction with other variables (e.g. Elkind, 1961a, 1964; Kofsky, 1966; Uzgiris, 1964) and computing in various ways an average age of attainment of various skills (e.g. Elkind, 1961c, 1962; Harter, 1979; Laurendeau & Pinard, 1962). Such techniques can confirm the tendencies of some attainments to develop at later ages than others, but unlike scalogram analysis, they say nothing about the invariance of such a sequence. Skill B may develop, on the average, at a later age than
skill A, but a significant number of individuals may still develop those skills in the order BA rather than AB. In a sense, scalogram and analysis of age as a variable give complementary information. Where a scalogram suggests an invariant sequence, we may still be interested in the average age of attainment of the steps in the sequence and the extent to which age is a significant variable. It is even possible, although perhaps unlikely, that skills A, B, and C would form an invariant sequence but there would be no main effect for age because the sequence occurs in a narrow time span. Even when scalogram analysis indicates that a sequence is not invariant, it may still be of interest to determine general developmental patterns, while remaining aware that some individuals will follow alternate sequences. Thus, techniques such as ANOVA and age of attainment data can be quite useful in analyzing developmental sequences, but such data should not be confused with the notion of invariant sequence, as some reviews (e.g. Pinard & Laurendeau, 1969) have done.

Using scalogram techniques, researchers have found invariant sequences for attainments in the sensorimotor period (Uzgiris & Hunt, 1975), for number concepts (Wohlwill, 1960), for conservation of quantity (Davol, Chittenden, Plante, & Tuzik, 1967; Uzgiris, 1964), for classification (Hooper, Sipple, Goldman, & Swinton, 1979), and for dream concepts (Kohlberg, 1966). Kofsky (1966)
found largely negative results in the area of classification skills, although certain pairs of items did seem to succeed each other in a nearly universal order, as assessed by interitem homogeneity coefficients. (Homogeneity analysis, Loevinger, 1947, is conceptually similar to scalogram analysis but allows for the testing of relationships between pairs of items as well as for the entire scale.) There were several differences between Kofsky's study and the study by Hooper et al., which found positive results in the same area. Kofsky included more tasks and her tasks tended to use a variety of materials whereas the materials in the Hooper et al. study were similar from task to task. It should also be noted that even Hooper et al. needed to eliminate some tasks in order to obtain an acceptable scale.

Dodwell (1960, 1961), working with number concepts, also obtained negative results regarding an invariant sequence. This study was more directly based on Piaget's work than was Wohlwill's (1960) study, which did find an invariant sequence in the development of number concepts. However, Dodwell (1961) attempted to form a single scale from tasks involving five different test subgroups (relation of perceived size to number, provoked correspondence, unprovoked correspondence, seriation, and cardination-ordination) rather than determining if the three stages within each test followed each other in a set order. The
latter approach would have been more in line with cognitive-developmental theory. Despite these exceptions, all in all the empirical results seem to support the claim that there are invariant sequences in the development of some cognitive abilities.

Nevertheless, certain qualifications are necessary. Some observations regarding measurement issues serve as a general warning in evaluating the results of any studies dealing with the temporal relationships of cognitive attainments. Flavell and Wohlwill (1969) highlighted the importance of the competence-automaton distinction. The competence model would refer to what the child knows or can do in some hypothetical ideal situation, whereas the automaton model (or the performance model, as it is more commonly known) would refer to what the child can actually do in a given situation. Flavell (1970) pointed out the importance of considering the relative difficulty of the tasks used to measure abilities, which is a "performance" factor, in drawing conclusions about any sequences. For example, if we find that the ability to perform task A invariably precedes the ability to perform task B, is this because in some meaningful way ability A and ability B constitute an invariant sequence of cognitive attainments, or merely because task A is an easier measure of ability A than task B is of ability B?
Braine (1959) criticized Piaget for not taking account of this issue and for apparently assuming that all tasks involving a particular ability would be equivalent. Braine proposed that the most meaningful measure of the attainment of a particular ability would be the earliest age at which that ability could be elicited. By extrapolation, it would seem that one meaningful way of defining an invariant sequence $AB$ would be that the earliest evidence of $A$ invariably precedes the earliest evidence of $B$.

Smedslund (1964) pointed out that in every concrete reasoning task, we can distinguish among percept (the specific stimuli used), goal object (e.g. length, quantity, etc.), and inference pattern (e.g. conservation, transitivity, etc.). He suggested that we define which factors are varying and which are being kept constant. He felt that the most interesting questions were relationships between inference patterns, with percepts and goal objects kept constant. In essence, keeping some factors constant is another way of equating tasks.

Flavell (1970) expressed doubt that this could be accomplished for all tasks. Nevertheless, it is worth noting that Kofsky (1966) did not find an invariant sequence for classification tasks using a variety of materials, whereas Hooper et al. (1979) did find such an invariant sequence using similar stimuli for each task. There
were other differences between the two studies which could have accounted for the different findings, but the need to find equivalent measures of tasks at different levels cannot be ignored.

A study of the developmental sequence of conservation skills (Uzgiris, 1964) illustrates how tasks can be equated, but raises other theoretical questions regarding invariant sequences. Let us ignore for now that facet of her study which involved the comparisons of responses using different materials (which is more relevant to the issue of structure than sequence) and consider responses to each material separately. As responses to each material were equally scalable, any material, say the commonly used plasticine, can serve as an example. Each material was subjected to three transformations. For example, the plasticine was changed from a ball, to a cylinder, to a longer cylinder, to three pieces. The child was considered to be a conservator if he or she evidenced conservation on all three transformations. The same material, the same transformations, and the same criteria of success were used in testing for conservation of substance, weight, and volume. Thus it seems reasonable to maintain that the tasks used to measure each separate ability were at equivalent levels of difficulty for their respective underlying abilities. Therefore, when these tasks were found to be scalable, it
was meaningful to maintain that these three forms of conservation follow each other in an invariant sequence.

The fact that classification skills, as assessed by Kofsky (1966) proved not to be scalable, whereas the sequence of conservation of substance, of weight, and of volume was scalable (Uzgiris, 1964) is somewhat ironic from the viewpoint of Piagetian theory. Kofsky's hierarchy of classification skills, derived explicitly from Inhelder and Piaget (1959/1964), represents precisely the kind of sequence that is basic to cognitive-developmental theory. On the other hand, the sequence studied by Uzgiris, although also based on Genevan research (Piaget & Inhelder, 1941/1974), is an example of a horizontal decalage. Although the concept of decalage has been incorporated into cognitive-developmental theory, the fact that such sequences are based on differences in content rather than changes of structure is certainly not basic to the theory. In fact, although Piaget and Inhelder (1941/1974) presented a persuasive, if post hoc, explanation of the sequence involved in the different forms of conservation, Piaget (1971, pp. 10-11) claimed that decalages in general are not predictable and remain a negative characteristic in the theory of stages.

For Piaget (1960/1971b, 1971, p. 9), the prediction of an invariant sequence of stages is based largely on the concept of equilibration. Certain earlier stages or
strategies are more probable at first, but once they fail to deal adequately with additional information, the next stage becomes logically necessary and inevitable, and so on through the entire sequence of stages, until a final equilibrium is reached. Thus for Piaget, invariance is more crucial in sequences of mental structures than in sequences based on content or decalage.

Of the studies discussed earlier, the Dudek and Dyer (1972) longitudinal study came closest to an actual test of Piagetian theory. The other studies of sequences involved decalages, extrapolations not necessarily basic to the theory, or other modifications of the theory. This is not to suggest that the finding of certain sequences is necessarily less interesting than others. Despite different roles in Piagetian theory, all invariant sequences (as well as sequences which are not necessarily invariant) are potentially important in understanding development. But it is important to be aware of the theoretical distinctions between different sequences.

Overall Structure

Those who have reviewed or summarized the empirical results on the question of overall structure, have generally concluded that the preponderance of evidence is against the strictly concurrent development of abilities hypothesized to be part of the same structure (Beilin,
A large number of studies have addressed this question, and only a few will be mentioned here. Some of the studies already discussed in support of the hypothesis of invariant sequence found negative results when they looked for synchronies between different sequences. Thus, Uzgiris and Hunt (1975) found low correlations between their different scales in the sensorimotor period. Similarly, Uzgiris (1964) found very high scalability for the sequence of three forms of conservation for four separate materials, but there was considerable variation across materials in that some subjects had achieved a certain level for one material but not for others. Similar variation of conservation abilities across materials was found by Lovell and Ogilvie (1960). Hooper et al. (1979), who found a scalable sequence in the development of classification skills, also found that the tasks represented a non-unitary structure from the viewpoint of factor analysis.

It was mentioned above that Dodwell's (1960, 1961) studies in the area of number concepts were among the minority that showed no evidence of invariant sequence. He also found no evidence for the synchronous development of different aspects of number conservation, all of which were hypothesized to be part of an operational understanding of numbers (Dodwell, 1960) or for the synchronous development of class inclusion and number concepts, which
had been hypothesized to develop together according to Piagetian theory (Dodwell, 1962). Some of the other studies that have not found consistently high correlations among Piagetian tasks hypothesized to be part of the same structure include Smedslund (1964), using a variety of concrete operational tasks, Beilin (1965) with conservation of length, number, and area, Shantz (1967) with three tasks involving multiplication of relations, and Lunzer (1960) with two different measures of conservation of volume and a measure of the understanding of continuity.

In one of the few clear exceptions to this pattern, Braine (1959) found an almost perfect correlation between levels of the ability to handle sequence relationships and understanding the concept of transitivity. Woodward (1959) found a strong pattern of concurrence between levels of sensorimotor intelligence and object concept development among severely retarded children. Less concurrence was found between stages of sensorimotor problem-solving and levels of circular reactions.

These empirical results on the question of developmental concurrences need to be examined from several points of view. First of all, there is the issue of the statistical analyses by which the results have been interpreted. As with the notion of invariant sequence, the notion of synchronous development of structurally related tasks translates most directly into a longitudinal design.
Over time, we would expect that as individuals attained new levels on one task, they would attain equivalent levels on a structurally-related task. However, Fischer (1980) reviewed a few studies which showed that the periodic re-testing involved in a longitudinal design increases the degree of consistency across tasks. Thus, it would seem that the measurement effect in a longitudinal study may be even more problematic for the structural hypothesis than for the invariant sequence hypothesis. At any rate, most of the studies that have addressed the question of synchrony and all of the ones discussed above have used a cross-sectional design.

Using a cross-sectional model, we should expect that at any moment in time all individuals would be at equivalent levels of any two abilities that are presumed to share the same underlying structure. This prediction can be readily tested using a contingency table. As Wohlwill (1973) pointed out, positive findings on the contingency table would not guarantee that all individuals actually arrived at the two equivalent levels at the same time, but if development on the two abilities is not synchronous, we should "catch" at least some individuals at two discrepant levels.

An advantage of a contingency analysis is that it can include not only a test of significance and an overall measure of the strength of the relationship between scores
on the two tasks, but also a qualitative analysis of the relationship between the two abilities when development is not entirely synchronous. For example, a contingency table can be used to evaluate whether one ability actually seems to precede the other. Several of the studies already discussed used contingency tables, although most did not include all three kinds of analysis (Braine, 1959; Dodwell, 1962; Lunzer, 1960; Smedslund, 1964; Woodward, 1959). Also, most studies used only a two by two table (dichotomizing passes and fails for each task) rather than a multilevel analysis. Woodward's (1959) study, which used five by five and five by four tables, showed the usefulness of such a design in comparing development through several stages on pairs of cognitive abilities.

The other statistical technique commonly used to evaluate synchronous development is correlation. Although it might first appear that the assumption of a continuous variable runs counter to the discrete nature of stages, even Piaget (1960/1971c, pp. 122-123) has commented that, to a certain extent, whether one gets continuity or discontinuity depends on the scale of measurement. Frequently, when a correlation coefficient has been used, a contingency analysis could have provided additional information. For example, Uzgiris (1964) correlated scores on conservation tasks across pairs of materials used in the tasks. Since
these scores corresponded to four scalogram patterns, she could have also set up four by four tables to analyze these results in terms of levels of conservation abilities. In such a case, a contingency analysis would generally provide a better test of the synchrony hypothesis. For example, if scores on one material were consistently one level above scores on another material, a perfect correlation could be obtained, while this would be an example of decalage rather than synchrony.

In some cases, correlation has no clear alternative in contingency analysis, other than arbitrarily dichotomizing the data. For example, Shantz (1967) correlated various scores, derived in different ways, all related to the concept of logical multiplication. Such a study would seem to bear on the relationships between such variables in a more general sense than a hypothesis of synchronous development of stages would imply, although this distinction may not always be clear. It would seem to be a fair conclusion that when data can be conceptualized in terms of levels or stages, a contingency analysis can give the clearest cross-sectional picture of the degree of synchrony between levels on pairs of abilities.

However, despite the fact that the statistical analyses have not always been optimal, it seems unlikely that the overall conclusion in regard to the issue of structure would be changed by improving the use of
statistics. In fact, the same general conclusion has usually been reached no matter which statistics have been used—that tasks presumed to have the same structural basis show some statistical relationship to each other, but not a strong enough relationship to justify the hypothesis of synchronous development.

In order to keep the concept of overall structures viable, some writers have suggested restricting its scope. For example, Wohlwill (1966) argued that the concept requires only consistency among various tasks employing the same operation or grouping, not all tasks in a particular stage or period. Furthermore, he pointed out that some relationships between groupings which had been hypothesized to develop concurrently turned out to be cases of invariant sequence (e.g. the relationship between transitivity and conservation of weight as studied by Lovell & Ogilvie, 1961). Thus, a model was suggested of concurrences within groupings and sequences between groupings.

Pinard and Laurendeau (1969), however, felt that Piagetian theory required synchrony across groupings, but only within a given concept (such as weight, distance, or class). While neither attempt to narrow the concept of structure fully accounts for all of the asynchronies that have been found, there does seem to be a pattern that the more limited the scope of a structural concept, the more synchrony is found. Of course, this tends to dilute the
strength of overall structures as a hypothetical construct.

Another attempt to deal with the issue of overall structure involves the competence-automaton distinction (Flavell & Wohlwill, 1969) that has already been discussed in the context of developmental sequences. As dealt with by Flavell (1970) in its simplest form, if we find that the ability to do task A precedes the ability to do task B rather than occurring concurrently as predicted, this may be because task A is an easier measure of ability A than task B is of ability B rather than because ability A really precedes ability B. However, if one were to take the most extreme position on overall structure— that all tasks involving a particular structure develop concurrently—this would make differences in difficulty between tasks involving the same structure meaningless. It is questionable whether anyone has ever clearly stated such an extreme position. Certainly Piaget (1956, p. 36; 1971, pp. 10-11) has incorporated the concept of horizontal decalage into his structural theory of stages. In any event, whether or not anyone has seriously considered the extreme form of the structural hypothesis, the data clearly refute it.

If we consider a more moderate hypothesis, allowing for tasks at different levels of difficulty within each cognitive structure, then we are confronted again with the problem of equating tasks for levels of difficulty. Thus, Flavell (1971) stated that a developmental
concurrence between two cognitive attainments hypothesized to be structurally related can mean at least four different things: (a) The two attainments emerge at the same time and reach functional maturity at the same time. (b) They only emerge at the same time. (c) They only reach functional maturity at the same time. (d) Development of one attainment is always ahead of the development of the other, but the two attainments develop at the same rate. Flavell felt, with some justification, that Piaget and close followers of his (e.g. Pinard & Laurendeau, 1969) would favor model a, although their writings are too ambiguous to be sure. Flavell also pointed out that Braine (1959) has supported using model b.

Similarly, Wohlwill (1973) defined four models of structural stage theory, which overlap Flavell's models, although they are not entirely parallel. Wohlwill's models are defined as follows:

Model I—Synchronous Development: All sequences within a structurally-defined stage develop concurrently, step by step. This was the only model that Wohlwill rejected as unrealistic and unsupported by the evidence.
Model IIA—Horizontal Decalages (convergent): There are systematic time lags between parallel sequences within a stage, but these time lags decrease.
Model IIB—Horizontal Decalages (divergent): The lags in development of one sequence behind another tend to increase.
Model III—Reciprocal interaction: One sequence at first lags behind another sequence, then overtakes it and in fact aids development in the other sequence.

Model IV—Disequilibrium-Stabilization: Developmental sequences that are related structurally go through transition periods in which there is increased variability across different tasks and which are followed by periods of increased consistency as stages are consolidated. This model was seen by Wohlwill as, in some ways, incorporating the other models rather than being distinct from them. He also seemed to favor Model IV overall.

Model IV is also similar to a four-phase model described earlier by Flavell and Wohlwill (1969). According to this model, in the initial phase, all tasks or items relating to a particular structure will be failed. The second phase is a transition period during which decalages and situational variables are relatively important, and thus there is a relative lack of consistency regarding different tasks that involve the same mental operation. The next phase involves increasing stabilization and consolidation of the overall structure, such that consistency among different tasks becomes more apparent. Finally, there is a phase in which the particular stage is firmly established and virtually all items involving a particular structure will be passed. Thus, as in Flavell's
(1971) model c, it is only when items reach their functional maturity that concurrences between the items would be expected.

These multi-phase models are appealing for several reasons. At a theoretical level, they seem to accord well with Piaget's (1956, p.35) comments about each stage having a period of preparation and a period of achievement, as well as some in-depth analyses of Piagetian theory (Massefat, 1963; Pinard & Laurendeau, 1969, especially section on "Consolidation"). This kind of model also emphasizes the developmental aspect of the concept of structure in cognitive-developmental theory, rather than viewing it as a static concept.

In addition to the theoretical advantages of a multi-phase model, there is also some empirical justification for it. It would be helpful to recall here that most of the studies reviewed earlier on the question of developmental synchronies did not find the consistently high correlations between tasks involving the same hypothesized cognitive operations that would have lent strong support to the concept of overall structure in its simplest form. Nevertheless, most of these studies did find positive correlations, and many of these correlations were statistically significant (Beilin, 1965; Hooper et al., 1979; Shantz, 1967; Uzgiris & Hunt, 1975).
When correlations have been computed separately for different age groups, patterns have at times been found which have been interpreted as favoring a disequilibrium-stabilization model (Flavell & Wohlwill, 1969; Wohlwill, 1973). The most direct test for such a model was a study by Nassefat (1963). His basic design was to test 9- to 13-year-old children on a variety of Piagetian tasks, grouped into three categories—concrete operations, intermediate, and formal operations. Then, for each category, correlations and scalogram analyses were computed at each age. Nassefat interpreted his results as favoring a four-phase model. In particular, he said they showed that for the age range of his sample, the concrete group of items had already achieved its period of maximum meaningful homogeneity and had moved into a period of trivial homogeneity (trivial in the sense that all items are passed); the intermediate items went from heterogeneity to homogeneity and finally toward trivial homogeneity as well; the formal items went from trivial homogeneity (in the sense that all items are failed) to maximum heterogeneity and only began to show some meaningful homogeneity.

Nassefat's study was cited prominently by Flavell and Wohlwill (1969; Wohlwill, 1973), and it clearly influenced their thinking. The phases they each refer to are essentially equivalent to each other, Flavell and Wohlwill's initial phase to Nassefat's first phase of trivial
homogeneity (when all items are failed), their transition phase to his phase of maximum heterogeneity, their phase of stabilization and consolidation to his phase of maximum meaningful homogeneity, and their final phase to his second phase of trivial homogeneity (when all items are passed).

Despite this apparent empirical support for the four-phase model, numerous questions regarding Nassefat's use of statistics and possible alternative interpretations of his data suggest the need for some caution. Because Nassefat's study is so pivotal to the disequilibrium-stabilization model and because the methodological questions raised by his study are important, these questions will be examined in some depth. Because part of the problem appears to relate to the use of a cross-sectional methodology to test a hypothesis with a time dimension, let us first specify what such a model would predict longitudinally for each individual. We would expect each individual to go through an initial phase of failing (i.e. scoring 0) on all items. Then he or she would score at level 1 on some items, while remaining at level 0 on other items. As stage 1 consolidates, we would expect this individual gradually to reach level 1 on all items.

If items are dichotomized as pass-fail (0, 1) this turns out to be a trivial hypothesis. However, if there are levels beyond level 1, then this model generates the testable hypothesis that level 1 will be achieved on all
items before the individual achieves level 2 on some items. Alternatively, in less deterministic form, as suggested by Wohlwill (1973), we can predict a period of increased consistency at level 1 before a transition period to level 2, which would involve relative inconsistency of levels among different items.

To test this model, Nassefat used scalogram analysis and contingency analysis, each computed separately by age. The use of scalogram in this way, figured separately by age and used to examine structure rather than sequence, is unique. Nassefat's assumption was that items associated with a particular stage (concrete operations, intermediate, or formal operations) should be scalable when that stage is consolidated, but not before. While the notion of scalability does refer to a kind of "homogeneity," as Nassefat calls it, and this may appear to be linked to notions like "consistency" and "consolidation," a closer look reveals a different underlying meaning. Finding that items are scalable at some ages but not at others does not imply that, for an individual, the consistency of his or her responses (in the sense of reflecting a similar level of thinking) is greater at those "scalable ages" than at "non-scalable ages." Nassefat's scalability findings could imply that individuals go through only certain elements of the sequence in the same order, and these elements are more
prominent at the scalable ages. However, there is a simpler explanation of Nassefat's findings. The scalable ages are those ages at which there is enough variability among the subjects to bring out the basic sequential nature of the tasks. The non-scalable ages, on the other hand, are those ages at which such large proportions of the children either fail all the tasks or pass all of them that the Guttman scale is not apparent. Nassefat's results seem to accord with this interpretation in that it is Green's index of consistency (which is highly dependent on marginal probabilities), rather than Guttman's reproducibility, which varies greatly from age to age. Also, as Nassefat himself pointed out, scalability is highest when the discrimination index for that group of items is highest. Thus, in measuring scalability by age groups, Nassefat appears to have found a useful way of locating those ages at which there is enough variability among subjects to make scalogram a meaningful technique, but he has not found a good measure of stage consolidation within subjects.

The other technique used by Nassefat, categorizing answers to each item along a continuum and then analyzing pairs of items with contingency tables, was potentially more relevant to the structural stage consolidation theory. However, Nassefat's use of a rank order correlation to analyze these data raises problems with this part of his
study as well. He settled on a rank order correlation after he ruled out the usual contingency coefficients because they do not take the ordinal relationships among response categories into account. He also rejected product-moment correlation because of its parametric and interval-scale assumptions. Wohlwill (1973) pointed out that Goodman and Kruskal's (1954) gamma would have been the ideal statistic for this situation. Wohlwill went on to question Nassefat's use of Kendall's tau because of the large number of ties in rankings that were obtained and because the probability level accepted by Nassefat (.10) was suspect.

An even more basic criticism is that a rank order correlation would be more sensitive to patterns among subjects at different ages than to patterns of responding to different items by each subject. Nassefat seemed to recognize part of this problem in that he noticed that when a large percentage of subjects fell within a single cell of a contingency table, the rank order correlation was low even if this cell represented the attainment of equivalent levels on the two items. To deal with this situation, Nassefat defined as a "privileged positive association" any relationship in which 2/3 of the population fell in a single cell representing equivalent levels on two items. He then added these privileged associations with significant rank order correlations to obtain the percentage of significant
interitem associations for each age group. Wohlwill questioned the meaningfulness of this hybrid measure of association. More fundamentally, it ignores the potentially common situation in which a significant tau is obtained despite marked deviations in the levels of the two items being correlated, when it is supposedly this consistency among items that is being measured. Once again, as with the scalogram analysis, relationships between subjects are being confounded with cognitive structures within subjects. An inspection of the actual contingency tables can help to clarify these relationships.

Although Nassefat's study was the most direct test of their hypothesis, Flavell and Wohlwill (1969; Wohlwill, 1973) also found support for their ideas in Uzgiris' (1964) conservation study. The correlations she reported of conservation abilities across different materials at different ages suggested that there are periods of relative inconsistency across materials followed by periods of relative consistency. Her use of statistics appears less questionable than Nassefat's. But an examination of the underlying contingency tables by grade level, to see if increases in correlations were due to increases in subjects showing corresponding levels of conservation would have been useful, especially since the differences in correlations was not great.
Uzgiris also reported the percentage of subjects showing "oscillating" responses (i.e. showing conservation for at least one but not all three transformations of a given material). This data was presented separately by grade levels and by kind of conservation—substance, weight, and volume. Oscillations for substance conservation decreased after second grade, at which point this achievement is consolidated. At the same point, oscillations for weight conservation increased, and then they decreased after fourth grade. This data does suggest patterns of consistency and inconsistency of responses over time, but only longitudinal data could answer such questions as whether individual subjects tend to show a period of responding consistently to both substance and weight conservation questions in between periods of inconsistency regarding substance and inconsistency regarding weight.

It should also be pointed out that by using age or grade level, Uzgiris, like Nassefat, used an approximation to developmental level that will vary to the extent that age or grade level is correlated with a particular stage. To improve on this approximation, Flavell and Wohlwill (1969) reanalyzed Uzgiris' data. Specifically, they applied their four-phase model to patterns of passes and fails of Uzgiris' subjects on substance and weight conservation tasks, then assigned the subjects to one of the
four phases, and finally analyzed the number of oscillations by phase rather than by age or grade. They found the expected pattern of an increase in oscillations during the transition phase and a decrease at the final phase. However, to some extent, this result was an artifact of their definitions of phases. As an oscillation on any material was defined as a failure by Uzgiris and the final phase was defined as success on all four materials for both substance and weight conservation, oscillations in the final phase were possible only on volume conservation tasks. Thus, it is not surprising that oscillations would be fewest in the final phase. Nevertheless, these results are not totally artifactual in that total possible oscillations decreases with each successive phase and yet the actual average number of oscillations increases from phase 1 to phase 2, as would be predicted by the model.

In summary, the following conclusions regarding the notion of overall structure seem warranted. The relationship between tasks hypothesized to share the same underlying structure is rarely strong enough or specific enough to support a model of strictly synchronous development. Content as well as structure seems to determine the likelihood that an individual will be able to succeed on a specific task. The more similar the content of two tasks, the more likely it is that equivalent or corresponding levels on each task will be attained concurrently.
Inasmuch as there generally appears to be a statistical relationship between structurally related tasks, it is meaningful for research studies to continue to assess the strength of such relationships as well as the qualitative nature of these relationships.

A model of stages which postulates periods of relative inconsistency among tasks followed by periods of stage consolidation also merits further investigation, although it has received less valid empirical support than some of its proponents have claimed. In particular, caution is necessary in interpreting cross-sectional data, as it is important to keep in mind that the model refers to within individual patterning of abilities across time. Scalogram analysis, while providing useful data for other purposes, does not provide data relevant to this model. In general, contingency table analysis is useful in research on this model. Any comparisons of overall measures of association should be supplemented by inspection of the contingency tables to verify that differences in degree of association are due to differences in the degree to which different tasks are at equivalent levels.

In many studies there may be measures particular to those data (such as the number of subjects showing oscillations in Uzgiris', 1964, study) that will provide a useful indication of degree of consistency of responses which can
be compared across groups. Whenever possible, one should consider defining the relevant cross-sectional groups in terms of developmental level rather than age, although care needs to be taken to insure that the definition of developmental level does not artificially determine the measure of consistency used as the dependent variable.

If enough levels for each task can be identified, consideration should be given to using methods outlined by Wohlwill (1973) for testing this model. This involves representing the subjects in a contingency table according to the levels achieved on each of two tasks. Then the variability of responses on each task is analyzed separately by levels achieved on the other task. For each level of task A, the average deviation from the modal category for task B is computed. The pattern of these variability scores for task B across levels of task A is then analyzed. A complementary analysis can be done for the variability of task A across levels of task B. For each task we can determine whether there are periods of increased and decreased consistency of responses.

In the long run, the question to be considered is whether the structural relationships are strong enough to justify the concept of stages or whether the data are better accounted for by a model that predicts relatively separate sequences in different content areas.
The Social and Affective Domain

Although Piaget (1956, 1960/1971a, 1971) has expressed doubts as to how far his theory of stages can be applied beyond its original grounding in logico-mathematical cognition, others have attempted to apply this model to the development of cognition regarding social and affective issues. One of the first and best known ventures in this area was Kohlberg's (1963, 1969) work on moral development from a cognitive-developmental viewpoint. Kohlberg's (1963) conceptualization was based, to some extent, on Piaget's (1932/1965) earlier work on moral development, but also differed from it in several important ways. Nevertheless, Kohlberg (1963, 1969) stayed very close to Piaget's (1960/1971a) definition of stages. Thus, it should be no surprise that much of the research involving Kohlberg's stages of moral development has focused on the issues of invariant sequence and structured wholeness.

Recognizing that a scalogram analysis was not appropriate because his stages were disjunctive rather than cumulative, Kohlberg's (1963) first attempt to find evidence for an invariant sequence was to show that intercorrelations among the six stages approximated a quasi-simplex. That is, when the subjects' scores for each stage were intercorrelated with their scores for all the other stages, correlations tended to be higher between adjacent stages than between stages farther apart.
Turiel (1966), in a training experiment, found that it was easier to facilitate movement to one stage above a child's present stage than to other higher or lower stages. He controlled for the possibility that the experimental procedures themselves altered the course of development in the expected direction by exposing subjects to arguments for moral stages below as well as above their own stage. Rest (1973) found that moral stages that subjects could comprehend, which unlike predominant level of thinking or preference are cumulative, formed a Guttman scale.

Some longitudinal data referred to by Kohlberg and Kramer (1969) also seemed to support the notion of a developmental sequence in moral stages, although the invariance of these stages was somewhat in doubt because of some temporary regression by middle-class college students. However, as Kohlberg and Kramer pointed out, this was not true cognitive regression in that the college students appeared still to be able to comprehend the higher stages which formerly characterized their thinking. Overall then, these studies do support the hypothesis of an invariant sequence in moral stages, especially when comprehension rather than typical thinking is the criterion.

In terms of structure, three kinds of claims have been made about moral development—that there is consistency among each subject's responses to different moral
dilemmas, that moral development as defined by Kohlberg is related to cognitive development as defined by Piaget, and that moral development is related to the underlying cognitive ability to take the role of another person. Kohlberg (1969) claimed that his data supported the premise of consistency within the realm of moral judgment. He reported that an average of 50% of a subject's moral judgments fell into a single stage and that correlations between moral levels based on responses to different stories ranged from .31 to .75, with a median of .51. While these data do indicate a certain degree of consistency in level of moral judgment, they fall short of the degree of synchrony that was usually sought in studies involving other aspects of cognitive development. In fact, it was just such moderate correlations that led to a questioning of the concept of overall structures in logico-mathematical cognition. If Kohlberg is willing to accept such data as supporting the concept of overall structures, this implies a loosening of that concept.

Turiel (1969) presented a model of moral stage structural unity similar to Wohlwill's (1973) disequilibrium-stabilization model. Turiel predicted that early and late in development there would be relatively little "stage mixture" or responses at different stages, while during an intermediate period there would be relatively
more stage mixture. He presented data from various studies, which he claimed supported the model. However, Wohlwill (1973) pointed out that the data are not quite as supportive as Turiel claimed, especially in that those at low levels of moral development were still showing much stage mixture even after reaching adulthood, whereas Turiel had predicted a decline in stage mixture even if higher moral levels were not attained.

Keasey (1975) discussed two studies of his linking Kohlberg's stages of moral development with Piaget's stages of cognitive development. He hypothesized that certain levels of cognitive development were necessary, but not sufficient, for particular levels of moral development. Supporting evidence was presented in that concrete operations was necessary for stage 2 moral reasoning and formal operations was necessary for stage 5 moral reasoning. However, as predicted, attaining a given level of cognitive development did not guarantee attainment of any particular level of moral development. Because synchronous development was not being predicted, correlations between moral development and cognitive development which varied from moderate to negligible, according to the age of the subjects, were seen as supportive rather than damaging to the hypothesized model. This kind of necessary-but-not-sufficient model will be discussed more fully.
later in this review.

On the other hand, Damon (1975) found more support for a traditional model of synchronous development between parallel levels of logical and moral reasoning than for a model which considers certain levels of logical reasoning necessary for certain levels of moral reasoning. Although Damon's correlations and contingency tables certainly did not indicate perfect synchrony, the evidence for some degree of parallel development was strong, and there was clear evidence that some subjects developed a particular level of moral reasoning without having reached the equivalent level of logical reasoning.

It should be noted that Keasey used standard Kohlberg moral dilemmas, whereas Damon used his own measure of "positive justice," which seems to be related to just one aspect of Kohlberg's stages. Also, Keasey studied an older age group than Damon did. These and other differences between the Keasey and Damon studies make it difficult to suggest an explanation for the different conclusions they reached. In general terms, the best overall conclusion appears to be that, while there is a positive, non-random relationship between Piagetian cognitive development and stages of moral judgment, the precise nature of this relationship has not been fully defined.

The relationship between moral judgment and role-taking abilities is of great theoretical interest.
Role-taking as a concept has a significant history in the cognitive-developmental literature. The recognition of the importance of role-taking skills, which include the ability to experience another person as like oneself, to understand that the other may not see things the way oneself does, and to perceive things from the other's point of view, is usually attributed by cognitive-developmentalists to the early writings of Baldwin (1906) and Mead (1934). For Piaget (e.g. 1947/1966), role-taking was also an important concept, representing decentering in the social sphere.

Like other cognitive-developmental concepts, role-taking skills have been analyzed from the viewpoints of invariant sequence and overall structure. Flavell (1968) studied role-taking extensively and found a regular developmental pattern for at least some tasks. Selman (1971b) identified four levels in conceptual role-taking, which were highly related to age. Selman and Byrne (1974) found a similar age-related sequence using a different instrument, and also found that the child's ability to understand questions relating to each of those four levels formed a Guttman scale. A very high correlation was found between two different dilemmas used to assess role-taking ability (Selman & Byrne, 1974) and totally different measures of role-taking skills were found to be related to each other, although these relationships were not as strong
(Selman, 1971a, 1971b). Decentering on Piagetian tasks and decentering on role-taking tasks have also been found to be correlated abilities (Feffer & Gourevitch, 1960).

Kohlberg (1969) put much emphasis on the development of role-taking skills as underlying the development of moral judgment. Selman (1971a) found that Kohlberg's levels of moral judgment were in fact related to levels of role-taking skills. Furthermore, in a one-year follow-up of subjects who scored low in both role-taking and moral judgment, it was found that no child attained conventional (stage 3) morality without also attaining reciprocal role-taking, but several children attained reciprocal role-taking without attaining conventional morality. On this basis, Selman argued against synchronous development of role-taking and moral judgment skills and in favor of the hypothesis that the ability to understand reciprocal role-taking was a necessary but not sufficient condition for the development of conventional moral thought. However, Selman failed to take account of the fact that his cross-sectional data showed a significant number of subjects who had attained conventional morality without reciprocal role-taking abilities. Of course, this fact would seem to argue against the necessary-but-not-sufficient hypothesis.

As moral judgment and role-taking were among the first areas within the social and affective domain to be extensively studied from a cognitive-developmental viewpoint,
they can be seen as representative of this kind of endeavor. It should be apparent from the studies referred to that the themes of invariant sequence and overall structure are just as prevalent in this domain as they are in other areas of cognitive-developmental research. Although the cognitive-developmental model does not always seem to fully account for the data, it remains viable enough to warrant further research from this perspective.

More specifically, as with research in the logical-mathematical realm, the invariant sequence hypothesis has received more straightforward empirical support, while the structural hypothesis has required a more flexible interpretation. Basically, there have been three approaches in the social-affective area to dealing with the empirical fact that strictly synchronous development appears to be the exception rather than the rule. One approach is to hypothesize that, while there may be more variability across the span of development than is implied by the notion of overall structure, there are certain points in development when stages stabilize and consistency increases. Another approach is the necessary-but-not-sufficient hypothesis, which suggests that we should expect a systematic decalage rather than synchrony between certain attainments which may appear logically parallel. A final approach is the acceptance of a weaker standard of relationship between achievements on different tasks as evidence
of an overall structure. It is important to note that in at least some studies these approaches have been incorporated as part of the hypotheses of the study rather than being added as post hoc explanations. In addition, these three approaches are by no means mutually exclusive.

As already discussed, the paper by Turiel (1969) is one of the clearest examples of the first approach (i.e. consolidation at certain points in development), and partial empirical support was obtained in the area of moral judgment. The other two approaches have been integrated into a model of overall cognitive development proposed by Selman (1976b). In this model, both structure and content are considered important in determining a person's response to a cognitive task. Furthermore, structure and content are seen as relative terms. The general logico-mathematical stages defined by Piaget are hypothesized to provide the underlying structure for all cognition. According to the model, these structures are then applied in two general content areas—physical and social. Piaget, of course, studied mainly physical cognition. Social role-taking is seen as providing the underlying structure for such social content areas as moral judgment and interpersonal relations.

The relative nature of the concepts of structure and content becomes clear when we see that social role-taking is a content area when viewed in relation to logico-mathematical cognition, but it is a structural parameter in
relation to moral judgments and interpersonal relations. Two kinds of relationships among abilities are predicted, then, by this model. Where one set of abilities is seen as structurally more basic than the other (e.g. logico-mathematical cognition vis-a-vis role-taking or role-taking vis-a-vis moral judgment) a necessary-but-not-sufficient relationship is hypothesized. For any two abilities in which A is seen as providing the underlying structure for B, on tasks involving equivalent levels of A and B, we should find some subjects who fail both A and B, some who pass A but not B, some who pass both A and B, but none who pass B and not A. Between parallel tasks (e.g. moral judgment and interpersonal relations) which share the same underlying structure, approximately synchronous development is hypothesized, but not in any absolute sense. Thus, it is not at all embarrassing to this model when correlations between parallel tasks are far from perfect, because the importance of content is explicitly recognized.

Selman's (1976b) evidence for his model was substantial, if not totally consistent. First of all, he presented a series of correlations and partial correlations which were intended to support his hierarchy of relationships. For example, inasmuch as social role-taking was supposed to mediate between logical reasoning and interpersonal relations, the correlation between the latter two variables was predicted to drop when role-taking was
partialed out, and it did from .69 to .17. On the other hand, also in accord with the model, the correlation between logical reasoning and role-taking did not drop significantly (from .85 to .70) when interpersonal reasoning was partialed out. Other patterns were more difficult to interpret. Selman predicted little drop in the role-taking-interpersonal relations correlation when logical reasoning was partialed out. He stated that the drop from .79 to .39 supported his claim, but this drop does appear substantial even though the latter correlation remains statistically significant. Other correlational findings, such as the drop in the correlation between role-taking and moral judgment, from .61 to .17, when interpersonal relations was partialed out, are hard to explain using this model.

However, although the results from the correlational patterns are less clear than Selman seemed to believe, his contingency tables showed surprisingly good support for the model. Very rarely did any subject attain a particular level of a task without also attaining an equivalent level of another task believed to be necessary for it. Furthermore, this kind of contingency table data is more vital to the support of the theory than the correlational data is. This strong empirical support seems especially surprising in that Selman himself recognized two important problems in this empirical test of his model. Firstly,
he was aware of the problem of equating the difficulty levels of his separate tasks, as was discussed earlier in this review.

The second problem is even more serious. The model postulates a necessary-but-not-sufficient relationship between general logico-mathematical structures and social role-taking (and in turn, with the other areas of social cognition). However, general logico-mathematical reasoning is a hypothetical construct, which cannot be measured directly. The Piagetian measures that Selman used have physical content. But there is nothing in Selman's theory to suggest that physical cognition is necessary for social cognition. Selman seemed to feel that the Piagetian measures were less content-bound than the social cognition measures, and hence he called them measures of "logico-physical" cognition. Nonetheless, he was aware that he was not testing the relationship between logico-mathematical reasoning and social reasoning directly. In fact, what he might have been demonstrating was a systematic decalage between supposedly parallel levels of physical and social cognition.

The Ability to Conceptualize Conflicting Emotions

Selman's (1976b, 1977) research is important here not only because of the theoretical model he presented, but also because the particular content area of interpersonal
relations, which was the focus of his studies, is of direct relevance to our concern with the ability to conceptualize conflicting emotions (ACCE). The content categories within the area of interpersonal relations concepts include understanding of individuals or persons, understanding of friendships or dyadic peer relationships, and understanding of peer group relationships. Within the category of persons or individuals are several subcategories, one of which (subjectivity) comes very close to our notion of ACCE.

Selman (1976b, pp. 162-163) described the following sequence for subjectivity according to his stages of interpersonal relations concepts: At stage 0, conflicting emotions are denied. At stage 1, simultaneous conflicting emotions are conceptualized, but not toward the same object. At stage 2, conflicting emotions are conceptualized toward the same object, but these feelings are isolated from each other, usually temporally. At stage 3, a single object can arouse truly mixed feelings. At stage 4, mixed feelings can generate a synthesis, such as melancholy feelings from a mixture of happiness and sadness.

There has been no direct empirical test of the invariance of this particular sequence. But overall stage scores for interpersonal relations, computed by averaging across 17 issues of which subjectivity would be one, were found either to remain the same or increase in a two-year longitudinal study (Selman, 1977). A significant degree of
structured wholeness in the concept of interpersonal relations has also been empirically demonstrated. In an interview involving four interpersonal dilemmas, correlations between .61 and .81 were obtained across the four dilemmas. In addition, only 2 out of 47 subjects had scores more than one stage apart on any two dilemmas. In each case, a subject received as an interpersonal relations score the highest stage of reasoning shown during the interview. In the same study, a correlation of .71 was obtained between stages in the Persons and Dyadic Relations domains (Selman, 1976b). In another study, which used average scores rather than highest stage shown, correlations between domains ranged from .73 to .87, and a factor analysis showed that 61% to 65% of the variance was accounted for by the first factor (Selman, 1977). However, in none of the studies reported have any of the subcategories (such as subjectivity, which is of prime interest here) been considered separately or in relation to other cognitive abilities.

While Selman incorporated the conceptualization of affect as part of an overall conceptualization of interpersonal relations, others have studied the development of children's understanding of affect as a separate dimension. A few studies have looked at the development of the understanding of single emotions in young children (Borke, 1971; Gilbert, 1971). But Harter (1977, 1979),
like Selman, has concentrated on the understanding of multiple emotions, while studying this as an independent dimension. Harter (1979) was concerned, first of all, with children's understanding of their own emotions, and saw the need for later study to establish empirically the relationship between this understanding and the understanding of others' emotions. On the other hand, Selman (1976b) seemed to be studying mainly children's understanding of the emotions of others, although he did not explicitly recognize this distinction. In fact, in the tradition of Mead (1934), he regarded all cognition about affective matters as originating in social interactions. Harter's research is more in line with Cowan's (1978) proposition that it is important to maintain a distinction between emotional dimensions and social dimensions.

Harter (1977) began her interest in this area with the observation that children at a mental health clinic found it difficult to express feelings that appear to be contradictory. For example, children often felt that they were "all dumb" with nothing smart about them, or that they only loved their parents with no part of them that disliked their parents. Harter felt that in addition to the usual psychodynamic explanations of such difficulties, there was also a developmentally related cognitive factor in these difficulties. In particular, Harter (1977, 1979) saw the shift from the egocentric, perceptually dominated, alogical
reasoning of the preoperational child to the decentered conceptual thought of the concrete operational child as crucial in the development of ACCE. She drew parallels between the inability of preoperational children to solve conservation problems or multiplication of classes matrices because they focus on only one physical or logical dimension at a time and the inability of young children to conceptualize conflicting emotions because they focus on only one emotional dimension at a time. Such parallels between the development of concrete operational groupings and a more logically integrated view of the self and one's feelings had been briefly discussed by Piaget (1964/1967).

Harter (1979) took the cognitive-developmental hypothesis that she had formulated in the clinic and tested it with a normal population. Implicitly, she saw the conceptualization of conflicting emotions as part of the broader issue of the conceptualization of multiple emotions (conflicting or not), and it was this broader issue that she studied. In interviewing children aged 3 to 18, she found the following sequence in children's abilities to deal with multiple feelings: (a) Only one feeling can be conceptualized. (b) Two feelings can be conceptualized, but only in temporal sequence. (c) Two feelings can be conceptualized as simultaneous. The average age at which step b was reached was between 6½ and 7½, while step c was generally reached at around age 9.
Furthermore, no child was able to conceptualize simultaneous emotions who could not also conceptualize sequential emotions.

Harter (1979) also noted that when children were able to express multiple emotions two important dimensions, in addition to the temporal dimension of sequential vs. simultaneous, which tended to differentiate their responses were whether the emotions were of the same or different valence and whether they were expressed toward the same or different objects. Valence refers to whether an emotion is considered positive or negative. Thus, happy and sad would be opposite valence emotions, whereas happy and proud or sad and jealous would be same valence emotions. If a child says that he is both happy and sad about getting a new bike because he wanted one but didn't like the particular one he got, both feelings are about the same object. However, if he is happy about the bike but sad about losing a toy, then the feelings are about different objects.

Because the dimensions of valence and object (unlike temporality) were identified post hoc and thus not systematically explored in the interview, there was incomplete data on children's ability to deal with these dimensions. In particular, for these dimensions, the data indicated the form that spontaneous answers are more likely to take, not whether children can conceptualize different kinds of affective situations if asked about them. Nevertheless,
Harter's data are interesting. First she examined combinations of the temporal and valence dimensions. When children were asked about sequential emotions, generally more of them talked about two emotions of opposite valence than of the same valence, and the children who talked about opposite valence emotions were younger than those who talked about same valence emotions. When asked about simultaneous emotions, more children gave answers involving same valence emotions, but there were no age differences. Similar patterns were found when combinations of temporal and object dimensions were considered. When asked about sequential feelings, more children referred to different objects of the emotions than to the same object, and children using different objects were younger than those using the same object. For simultaneous emotions, generally more children applied the emotions to the same object, but there was no significant age difference.

Probably because of the small numbers in each category, Harter made no systematic tabulation of the sequence of categories involving the eight combinations of the temporal, valence, and object dimensions. However, based on a global inspection of the data and some intuitive hunches, she tentatively placed them in the following order: sequential, different valence, different objects; sequential, different valence, same object; sequential, same valence, different objects; sequential, same valence, same
object; simultaneous, same valence, same object; simultaneous, same valence, different objects; simultaneous, different valence, different objects; simultaneous, different valence, same object.

In discussing her results and suggestions for future research, Harter (1979) referred to the usual cognitive-developmental issues of invariant sequence and overall structure. Although her data pointed as clearly as cross-sectional data ever can to an invariant sequence of conceptualizing sequential emotions before simultaneous emotions, her lack of systematic data left open the question of invariance regarding the larger sequence involving combinations of valence, object, and temporality. She recommended including these dimensions in future studies and subjecting the results to scalogram analysis.

Regarding the issue of structure vs. content within the context of the ability to conceptualize multiple emotions, Harter defined structure as involving the various dimensions already discussed, while she saw content as referring to the particular emotions involved. Specifically, the interviewer inquired about four basic emotions—happy, sad, mad, and scared. The child then chose which of any other emotions to discuss in either a sequential or a simultaneous relationship with the initially presented emotion. Harter's data were analyzed separately for each of the four basic emotions presented by the interviewer.
While all four emotions seemed to follow the same pattern regarding the appearance of a sequential conceptualization before a simultaneous conceptualization, a few differences emerged when the other dimensions were considered as well. For example, when mad was the presented emotion and the children were asked what feeling they could have after feeling mad, there was no significant difference between the number of children giving same valence and different valence answers. Furthermore, the subjects giving different valence answers were older, which was the opposite of the age trend for the other three emotions. Harter's attitude toward such data was similar to Selman's (1976b) position that we should expect both structure and content to be significant in determining developmental levels. Actually Harter went one step further than Selman in accepting the possibility of a difference in the pattern of a developmental sequence rather than just a decalage in the rates of two similar sequences. In any event, for the most part, the developmental sequences for each of the four emotions appear to be the same, but more research is needed in this area.

Another structural relationship that Harter (1979) alluded to in her discussion was the hypothesized connection between the ability to conceptualize multiple emotions and Piagetian concrete operational thought. She presented no data on this matter, but recommended this as a fruitful
area for further research.

It has already been mentioned that although Harter's (1977) original concern was with the ability to conceptualize conflicting emotions, the study just described (Harter, 1979) dealt with the broader issue of conceptualizing multiple emotions. Nevertheless, if we define conflicting emotions as multiple emotions of opposite valence, the relationship between these concepts becomes clear. If we include only those conceptualizations involving opposite valence emotions, Harter's (1979) hypothesized sequence becomes: sequential, different objects; sequential, same object; simultaneous, different objects; simultaneous, same object. It is interesting to compare this with the sequence that Selman (1976b) proposed for subjectivity. Using Harter's categories, Selman's stages 1, 2, and 3 become: simultaneous, different objects; sequential, same object; and simultaneous, same object. Thus, the order of simultaneous, different objects and sequential, same object are reversed in the two hypothesized progressions.

Nevertheless, despite this discrepancy and despite some differences in overall conceptualization between Selman and Harter, there is much that the two of them agree on. Specifically, they both conceptualize ACCE within a cognitive-developmental framework. Thus, they expect the development of ACCE to follow an invariant
sequence, to show at least some degree of structured wholeness independent of content, and to be related to the development of Piagetian logico-mathematical abilities. It seems appropriate to test these hypotheses more rigorously.
METHODS

Subjects

A total of 60 subjects, all from the same middle-to upper-middle-class school district in the Columbus metropolitan area, participated in the study. There were 20 subjects in each age group—5-6, 7-8, 9-10, and 11-12. Each age group defined an inclusive age range (e.g. the 5-6 group could include children from 5 years 0 months through 6 years 11 months). Within each of the first three age levels, 10 of the subjects were boys and 10 were girls. It was not possible to achieve this balanced distribution in the 11-12-year-old age group, which contained 6 boys and 14 girls.

Volunteers were solicited from a single elementary school and the seventh grade of a junior high school, and subjects were chosen from among the volunteers so as to maximize the age and sex balance. Of those who volunteered, 17 children were not included in the study. Group intelligence test results (on the Short Form Test of Academic Aptitude, derived from the California Test of Mental Maturity) were available for most subjects in the third grade or above. These results indicated an average IQ of 110,
suggesting that this sample was of above average intelligence.

ACCE Interview

The ACCE interview used in this study was an adaptation of the interview developed by Harter (1979) to study the development of the ability to conceptualize multiple emotions. Her format was changed in order to inquire only about opposite valence (conflicting) emotions and to inquire specifically about feelings both toward a single object and toward two different objects. Another change involved varying the order of presentation of the questions. In Harter's interviews, all children were questioned first regarding sequential emotions and then regarding simultaneous emotions. Thus, it could be argued that her results indicating that simultaneous conceptualizations are more difficult than sequential conceptualizations were due to greater fatigue or lack of interest later in the interview. Therefore, in the present study, in each age by sex group, half of the children were asked first about sequential emotions and half were asked first about simultaneous emotions.

Some additional changes were introduced after a short pilot study. Harter had asked questions such as, "What feeling could you have after you feel happy?" Then the children were asked what would make them feel that way.
Thus, her question did not specifically ask the children to put the two feelings together sequentially. The older children generally indicated in some way that there was a relationship between the two feelings. But the younger children often seemed to answer this kind of question as if they were answering a question regarding a single emotion, rather than multiple, sequential emotions. Thus, this question was changed to, "Tell me how you could feel happy and then have some bad feeling after that," in that way requiring them to mention the two feelings together.

Responses in the pilot study to questions about sequential feelings toward the same object fell into two very different categories. Most of the children gave examples in which the situation itself changed (e.g. "I was sad because it was raining and we couldn't go to the baseball game. Then the sun came out, and I felt happy"). A much smaller number of children gave responses in which the situation did not change, but the child later saw the situation in a different light (e.g. "I was sad that my grandfather died. Then I thought about it and realized that he was in heaven, and I felt better"). Since the

---

1 This and similar illustrative quotations are similar to statements made by children in actual interviews, but are not the specific words used by these children. Only in the Examples in the Appendix are such quotations verbatim.
cognitive level of these two kinds of responses intuitively seemed quite different, it was decided to ask separate questions to elicit each kind of response. These two categories of response were called sequential, same object, external change and sequential, same object, internal change, respectively. Harter had hypothesized that with sequential responses, different object responses would precede same object responses. Based on the pilot study results and this writer's intuition, it was hypothesized that different object responses would be intermediate between the two different types of same object responses.

Overall, then, there were five categories of response, which in the hypothesized developmental order are as follows: sequential, same object, external change; sequential, different objects; sequential, same object, internal change; simultaneous, different objects; simultaneous, same object. For brevity, these five categories frequently will be referred to as A, B, C, D, E, respectively. Questions to elicit each of these kinds of response were asked regarding four primary emotions—happy, sad, mad, and scared.

Preceding the part of the interview that was scored, the children were asked to name as many feelings as they could think of and then to give an example of what would make them feel that way. Any of nine common emotions (happy, sad, mad, scared, loving, jealous, ashamed, proud,
and nervous) which were omitted by a child were supplied to him or her, and the child was then asked for a situation in which such an emotion would be experienced. Harter had originally used this part of the interview to obtain data on the number of emotions and which particular emotions children at different ages make use of. In the present study, only the responses to the first four emotions were actually used, in that the children were reminded of those responses later in the interview. Nevertheless, this part of the interview was retained in full as a way of helping the children begin thinking about a variety of emotions. In addition, the children were asked to label each of the nine emotions as a "good" feeling or a "bad" feeling, as a way of introducing the concept of valence in a way that even the youngest children could understand.

Following this warm-up part of the interview, each child was asked either about sequential or about simultaneous conflicting emotions. For a child asked about sequential feelings first, the interviewer would say at this point, "You told me that one thing which made you happy was ..." (and the situation mentioned by the child in preliminary part of the interview would be repeated) "and you said that being happy was a good feeling. Tell me how you could be happy and then have some bad feeling after you're happy." From there the interview proceeded according to what kind of situation the child mentioned. If the
example referred to a single object with an external change, the interviewer would proceed, "It sounds like both of those feelings were about the same thing, about X, and that something happened in between to change your feelings. Now think of how you could be happy about something, and then nothing happens, but maybe you just think about the situation some more, and then you have a bad feeling about it." Further explanation was given, using the child's examples, if this was not clear. Finally, the child would be asked, "Now tell me how you could be happy about one thing and then have a bad feeling about something completely different." This same format was repeated for each of the other three emotions, with the child being asked how he or she could have a good feeling after feeling sad or mad or scared.

The questions regarding simultaneous conflicting emotions proceeded in a similar manner. The interviewer would say, "Before we talked about how you could have one feeling, then have a different feeling after it. Now, I'm going to ask you about having two feelings at the very same time. You told me before that one thing that made you happy was X. Now tell me how you could have a bad feeling at the very same time that you were feeling happy." If this time the child gave a response involving two different objects, the interviewer would continue, "That sounded like those two feelings were about two different
things. Now think of a situation where the happy feeling and the bad feeling are both about the same thing." Again this procedure was repeated for each of the three remaining emotions.

Note that although the presentation of the simultaneous questions first or the sequential questions first was set by the interviewer, the order of the questions within each of these broad categories varied according to the answers spontaneously given by the subjects. In each case, however, all five categories of ACCE were asked about for all four emotions. In general, the same order of emotions—happy, sad, mad, scared—was followed for all subjects. However, if the child spontaneously began to talk about a different emotion, the order was varied to move with the child's train of thought. Also, in discussing conflicting simultaneous feelings with the emotion happy, the child most often chose sad as the other feeling. At times, to avoid confusion, the interviewer would ask next about mad rather than sad.

Frequently, responses were ambiguous as to the category they belonged in. There was uncertainty about whether responses dealt with simultaneous feelings or sequential feelings and whether they pertained to a single object or two objects. In such cases, the children were asked to explain how they conceptualized the answer. Other responses were not ambiguous, but clearly did not fit the
category asked about. The interviewer would then give the child feedback as to what category he thought the child's answer fit, and then the child would be given at least one more chance to think of an appropriate example.

It should be noted that this questioning also differed from the technique used by Harter, who apparently accepted the child's initial response. The motivation behind these questions was the same as the Piagetian "methode clinique": to go beyond the face value of the child's initial responses, to understand the thinking behind the responses, and to determine if the child is capable of thinking in ways he or she may not at first show. Nevertheless, the kind of feedback given and the specificity of the questions in this interview probably went beyond the usual methode clinique, which tends to be less directive. Despite the problems inherent in this more directive method, it was felt necessary both to help the children understand the task and to help the interviewer understand the children's ambiguous responses. It also seemed that when children could not conceptualize in a given way, their answers betrayed this fact no matter how leading the questions were.

Each child was scored as passing or failing each of 20 items (5 categories by 4 emotions) depending on whether or not he or she showed evidence of such a conceptualization at some point in the interview. Each response was
scored according to the category it best fit, unless it did not represent a combination of two opposite valence feelings at all. Thus, the subjects could be credited with passing a particular item when they were being asked about a different item. When subjects passed a particular item more than once during the interview, they were still only credited with a single pass. That is, all items were scored 0 or 1, with no scores above 1. Most of the analyses involved the use of these separate dichotomous scores or scores representing developmental levels, which were based on these item scores, as will be explained in the Results section of this study. For some correlations, total scores were used, in which case the 20 item scores were simply summed.

Despite the extensive questioning, some responses remained ambiguous. The Appendix presents rules for scoring responses and verbatim sections of interview protocols, followed by the scoring of the responses and an explanation of the scoring. This will give the interested reader a clearer picture of the interview technique as well as the scoring system.

Ten of the 80 protocols were chosen at random and scored by an independent rater to establish the reliability of the scoring system. Before scoring the 10 protocols, this rater was given an explanation of the scoring rules
and was then trained on a single protocol, after which differences with the primary rater were reconciled. This training protocol was not included in the 10 used to assess reliability. The percentage of items (out of 20) on which the two raters agreed was computed for each of the 10 protocols. Agreement on individual protocols ranged from 70% to 100%, with an average agreement of 94.5%. The product-moment correlation between the total scores of the two raters was .98.

Measures of Logico-physical Reasoning

The children's understanding of conservation was assessed according to the general methods first developed by Piaget and Inhelder (1941/1974), but a more standardized approach was used, as others have done (Elkind, 1961b; Uzgiris, 1964). Each child was shown two balls of clay and asked if they had the same amount of clay in them. Children who did not think so were encouraged to find some way of making the balls the same. Once the child was convinced the two balls had the same amount of clay, the experimenter rolled one into a "sausage," and the child was asked if the two items had the same amount of clay or if one had more, and why. The sausage was then rolled back into a ball, and the child was again given a chance to equalize the balls, if necessary. Then the experimenter
flattened one ball into a "pancake" and the same questions were asked. Finally, the same procedure was followed with the one ball being divided into four smaller balls. Following this procedure for conservation of substance, the same procedure was repeated, with the experimenter this time asking equivalent questions regarding weight. For conservation of volume, a beaker of water was used with the clay. The child drew a line on the beaker indicating the original water level, and then drew another line indicating the water level after a ball of clay was dropped in. Then the same ball of clay was transformed successively into a sausage, a pancake, and four balls. The children were asked whether the water would rise, fall, or stay where it was. When they answered that it would rise, as most children did, they were asked how much it would rise and why. Finally, they were asked if they thought the sausage, pancake, or four balls would take up the same amount of room or space in the water as the one ball did.

The scoring of the conservation responses was similar to that used by Uzgiris (1964). Subjects were considered to pass a particular mode of conservation (substance, weight, or volume) only if they showed conservation responses on all three transformations of the clay. For conservation of substance and weight, the children's explanations proved not to be crucial, especially because of the stringent criterion for success. Thus, a particular
response was scored as evidence of these two forms of conservation if a child identified the amounts of clay or their weights as the same. On the conservation of volume task, many children predicted that the water would rise the same amount without any apparent concept of volume being applied. Thus, in addition to predicting the water level correctly, the child needed either spontaneously to explain the rise in water as attributable to equality of volume or to answer that the two objects took up the same amount of room in the water when asked in order to get credit for conservation of volume. In addition to the separate pass-fail scores for conservation of substance, weight, and volume, a total conservation score (which could vary from 0 to 9) was computed by summing the conservation responses for all three modes of conservation.

Unlike the rather clear-cut decision involving the choice of conservation tasks, the choice of classification tasks leaves more room for discretion. The main Piagetian work in this area (Inhelder & Piaget, 1959/1964) discussed a variety of skills and techniques, and researchers since then have devised various semi-standardized adaptations and extrapolations of those techniques. The present study relied heavily on the recent work of Hooper et al. (1979), who found that a number of classification tasks could be grouped into six ability levels. According to their results, ability grouping I was too easy and ability
grouping VI too difficult for the age range of this study. Thus, one task was chosen from each of levels II through V and adapted only slightly for this study. The main changes involved the use of somewhat larger stimuli and pieces of cardboard rather than blocks. The tasks used are as follows:

**Matrices** (grouping II)

This was the task that was modified the most from Hooper et al. In this study, only one kind of matrix task was used, what Hooper et al. referred to as a cross-classification matrix, as distinguished from matrices that involved seriation. Since they used only one matrix of this sort (the first one described below), two similar matrices were devised for this study. Furthermore, Hooper et al. used three tasks (replacement, reproduction, and transposition) with each matrix, and this study used only reproduction.

For each matrix, the experimenter asked the subject to watch while he arranged the stimuli on a white piece of linoleum, which had been clearly marked into boxes, one stimulus to each box. This arrangement formed a matrix according to two class attributes of the stimuli. After the child had a chance to look at the resulting arrangement, the stimuli were removed and the child was asked to put the pieces of cardboard back in the same places. The first matrix involved six triangles of three colors (blue, yellow, and red) and two sizes. The second matrix involved six red...
pieces of cardboard of three shapes (square, circle, and triangle) and two sizes. The last matrix involved nine pieces of cardboard of three shapes (square, circle, and triangle) and three colors (blue, yellow, and red). Each of the three subitems was considered passed if an acceptable matrix was produced, even if it was not the same matrix shown by the experimenter. The overall item was considered passed only if all three matrices were reproduced.

**Some-All (grouping III)**

Three small red triangles, two small blue triangles, and four small blue squares were placed before the subject in a line, oriented in the same direction, with the reds mixed in with the blues and the triangles with the squares. The child was asked to label the two colors and the two shapes. Then the following four questions were asked: "Look at all the red ones.... Are all the red ones triangles?" "Look at all the triangles.... Are all the triangles red?" "Look at all the squares.... Are all the squares blue?" "Look at all the blue ones.... Are all the blue ones squares?" The child was given one point for each correct answer and needed to get all four questions correct to pass this item.

**Dichotomies (grouping IV)**

Three small red circles, three small blue circles, three small red squares, three small blue squares, two large red
circles, three large blue circles, three large red squares, and two large blue squares were placed before the child in a scrambled fashion. Two small boxes were placed on either side of the stimuli, and the subject was asked to place the pieces of cardboard in the two boxes so that all of one kind went in one box and all of another kind went in the other box. The child was then asked to explain how the objects were sorted. After this was done, the stimuli were removed and rescrambled, and the child was asked to do this again in a different way. Each subject was given at least three sorting trials. If the child achieved at least two out of the three acceptable sortings (color, shape, size) in the first three tries, he or she was given a fourth attempt. The subject was given a point for each acceptable sort, and the overall task was considered passed only if all three sorting criteria were used.

**Class inclusion** (grouping V)

All stimuli were presented to the child in a linear manner, similar to that used in the Some-All task. After the presentation of each set of stimuli, the child was asked to label each attribute and to count the members of each class. In the first subitem, the stimuli were three small red triangles and two small blue triangles, and the subject was asked, "Are there more triangles or more red ones?" The next set of stimuli included three small yellow circles and two small blue circles, and the question was, "Are there
more blue ones or more circles?" Finally, three small red triangles, two small blue triangles, and four small blue squares were presented, and three questions were asked: "Are there more triangles or more red ones?" "Are there more blue ones or more squares?" "Are there more blue ones or more triangles?" If at any point the use of the word "ones" was confusing to the child, the words "pieces of cardboard" were substituted. Potential scores on this task ranged from 0 to 5, and a score of 5 was considered passing.

As with the other measures used in this study, the classification responses were analyzed in two ways. Primary consideration was given to the passes and fails on the four individual items. In addition, a total classification score was computed by summing the correct responses on all subitems. This total score had a potential range of 0 to 15.

Overall Procedure

The subjects were all seen on a one-to-one basis in a room at their schools during regular school hours. The subjects were given an introduction to the tasks which consisted of telling them that the purpose of the session was to find out how different children thought about these questions. They were also told that their answers would in no way affect their grades or anything else that happened to them at school. The subjects were then given
the warm-up and the first actual section of the ACCE interview, either the sequential questions or the simultaneous questions. The conservation tasks were interspersed between the two main sections of the ACCE interview as a way of relieving the monotony of the interview questions. After the conservation tasks, the remainder of the ACCE interview was administered. Finally, the classification tasks were presented in the following order: dichotomies, Some-All, class inclusion, matrices.

Most of the children completed all of the tasks in a single session, although the sessions were occasionally interrupted by recess, lunch, etc. Some of the children in the youngest age group were tested in two sessions because they became tired and could not concentrate any longer. The first session usually contained the first part of the ACCE interview and the conservation tasks, while the second session included the second part of the ACCE interview and the classification tasks. Generally, the sessions were separated by only a day or two. In a few cases, some of the older children seemed to tire after the first half of the session, but whenever they were given the choice of completing the task at another time, they chose to continue and finish in one session.

All of the interviews and other assessment procedures were administered and scored by the present writer. The interviews and the conservation tasks were tape-recorded
for scoring after the sessions, while the classification
tasks were scored immediately during the sessions. The
total testing time for each child was approximately 1 to
1½ hours.
RESULTS

Overview

Responses to the ACCE interview were considered separately for each emotion, but it was also felt desirable to define an overall measure of ACCE. An individual was considered to pass this overall measure for a particular response category if he or she passed that category on any three out of four emotions. Tables 1 through 5 summarize the percentage of subjects showing ACCE in five categories for this overall measure and for each emotion separately.

Looking first at the overall scores (Table 1), we see that the order of difficulty of the items tends to be ABEDC (with very little difference between D and C), rather than the predicted order of ABCDE. This means that conceptualizing conflicting emotions simultaneously toward the same object is apparently not the most difficult of the tasks, but is easier than conceptualizing conflicting emotions simultaneously toward different objects or sequentially toward the same object when there is no external change.
### Table 1
**Percentage of Subjects Passing Each Category of Response for Overall Scores**

<table>
<thead>
<tr>
<th>Age Group</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>5-6</td>
<td>50</td>
<td>30</td>
<td>0</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>7-8</td>
<td>80</td>
<td>55</td>
<td>15</td>
<td>30</td>
<td>35</td>
</tr>
<tr>
<td>9-10</td>
<td>90</td>
<td>90</td>
<td>55</td>
<td>50</td>
<td>65</td>
</tr>
<tr>
<td>11-12</td>
<td>95</td>
<td>80</td>
<td>70</td>
<td>65</td>
<td>75</td>
</tr>
<tr>
<td>All S's</td>
<td>79</td>
<td>64</td>
<td>35</td>
<td>36</td>
<td>15</td>
</tr>
</tbody>
</table>

### Table 2
**Percentage of Subjects Passing Each Category of Response for Happy**

<table>
<thead>
<tr>
<th>Age Group</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>5-6</td>
<td>70</td>
<td>35</td>
<td>5</td>
<td>20</td>
<td>25</td>
</tr>
<tr>
<td>7-8</td>
<td>95</td>
<td>60</td>
<td>30</td>
<td>45</td>
<td>40</td>
</tr>
<tr>
<td>9-10</td>
<td>100</td>
<td>85</td>
<td>70</td>
<td>75</td>
<td>75</td>
</tr>
<tr>
<td>11-12</td>
<td>95</td>
<td>100</td>
<td>70</td>
<td>80</td>
<td>80</td>
</tr>
<tr>
<td>All S's</td>
<td>90</td>
<td>70</td>
<td>44</td>
<td>55</td>
<td>55</td>
</tr>
</tbody>
</table>

### Table 3
**Percentage of Subjects Passing Each Category of Response for Sad**

<table>
<thead>
<tr>
<th>Age Group</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>5-6</td>
<td>70</td>
<td>40</td>
<td>0</td>
<td>20</td>
<td>5</td>
</tr>
<tr>
<td>7-8</td>
<td>80</td>
<td>70</td>
<td>20</td>
<td>45</td>
<td>55</td>
</tr>
<tr>
<td>9-10</td>
<td>85</td>
<td>95</td>
<td>60</td>
<td>60</td>
<td>65</td>
</tr>
<tr>
<td>11-12</td>
<td>95</td>
<td>75</td>
<td>80</td>
<td>60</td>
<td>65</td>
</tr>
<tr>
<td>All S's</td>
<td>83</td>
<td>70</td>
<td>40</td>
<td>46</td>
<td>48</td>
</tr>
</tbody>
</table>
### Table 4
Percentage of Subjects Passing Each Category of Response for Mad

<table>
<thead>
<tr>
<th>Age Group</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>5-6</td>
<td>40</td>
<td>35</td>
<td>15</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>7-8</td>
<td>70</td>
<td>65</td>
<td>15</td>
<td>35</td>
<td>35</td>
</tr>
<tr>
<td>9-10</td>
<td>90</td>
<td>95</td>
<td>60</td>
<td>60</td>
<td>75</td>
</tr>
<tr>
<td>11-12</td>
<td>90</td>
<td>90</td>
<td>70</td>
<td>75</td>
<td>65</td>
</tr>
<tr>
<td>All S's</td>
<td>73</td>
<td>71</td>
<td>40</td>
<td>44</td>
<td>45</td>
</tr>
</tbody>
</table>

### Table 5
Percentage of Subjects Passing Each Category of Response for Scared

<table>
<thead>
<tr>
<th>Age Group</th>
<th>A</th>
<th>E</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>5-6</td>
<td>50</td>
<td>45</td>
<td>5</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>7-8</td>
<td>75</td>
<td>50</td>
<td>40</td>
<td>35</td>
<td>50</td>
</tr>
<tr>
<td>9-10</td>
<td>90</td>
<td>90</td>
<td>80</td>
<td>60</td>
<td>70</td>
</tr>
<tr>
<td>11-12</td>
<td>90</td>
<td>65</td>
<td>70</td>
<td>65</td>
<td>90</td>
</tr>
<tr>
<td>All S's</td>
<td>76</td>
<td>63</td>
<td>49</td>
<td>40</td>
<td>54</td>
</tr>
</tbody>
</table>
We also note that, for each category, there is a trend for increased success with increased age, and that at each age level the pattern of relative difficulties of the categories remains fairly consistent. If we define age of attainment as the first age group at which at least 50% of the subjects pass an item, then we note that 5-6-year-olds can deal with sequential conflicting emotions toward the same object when there is an external change, 7-8-year-olds with sequential conflicting emotions toward different objects, while the other three categories are all attained at the age of 9-10.

An examination of Tables 2 through 5 reveals that the pattern for each separate emotion is similar to that for the overall scores. For the total subject group, scared is the only emotion for which there is a reversal of the ABEDC order. Even this one reversal may not be significant in that only the order of categories C and D were interchanged, and the percentages for these two categories are very close in the overall scores. Although the order of category difficulty is similar for each emotion, it does appear that the emotions are not equally difficult for the children to deal with. Generally, the subjects did best when the emotion presented to them was happy, next best with sad, then mad, then scared, although this order did not hold for all categories. It should also be noted that this order was generally the order in which the emotions
were presented to the children.

The effects of the different variables were also tested using an analysis of variance in which categories and emotions were within group measures and age and order of presentation (sequential first or simultaneous first) were between group measures. This analysis of variance is summarized in Table 6. Significant main effects were found for age, category, and emotion. No main effect was found for order of presentation, nor were any of the interactions involving this variable found to be significant. A significant interaction was found for age by category. This interaction is illustrated in Figure 1. As that graph shows, the main factor in the age by category interaction is a gradual leveling off of category effects, meaning that at successive ages there is less of a difference in the degree of difficulty of the various response categories. The emotion by category interaction is not statistically significant, suggesting that the slight difference in the order of the categories for the emotion scared may not be reliable.

The effect of sex was tested separately. The mean total scores for girls and boys were very close (11.35 and 11.77, respectively), and this difference was not statistically significant, t(78) = 0.29, p > .05. Thus, all further analyses were done with boys and girls combined.
Table 6
Analysis of Variance Summary

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>78.69</td>
<td>3</td>
<td>25.74(^b)</td>
</tr>
<tr>
<td>Order (of presentation)</td>
<td>0.02</td>
<td>1</td>
<td>0.02</td>
</tr>
<tr>
<td>Age X Order</td>
<td>2.78</td>
<td>3</td>
<td>0.91</td>
</tr>
<tr>
<td>Error(^a)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Emotions</td>
<td>1.48</td>
<td>3</td>
<td>4.17(^c)</td>
</tr>
<tr>
<td>Emotions X Age</td>
<td>0.88</td>
<td>9</td>
<td>0.83</td>
</tr>
<tr>
<td>Emotions X Order</td>
<td>0.63</td>
<td>3</td>
<td>1.79</td>
</tr>
<tr>
<td>Emotions X Age X Order</td>
<td>0.90</td>
<td>9</td>
<td>0.85</td>
</tr>
<tr>
<td>Error(^a)</td>
<td>25.49</td>
<td>216</td>
<td></td>
</tr>
<tr>
<td>Categories</td>
<td>32.55</td>
<td>4</td>
<td>37.70(^b)</td>
</tr>
<tr>
<td>Categories X Age</td>
<td>5.36</td>
<td>12</td>
<td>2.07(^d)</td>
</tr>
<tr>
<td>Categories X Order</td>
<td>1.36</td>
<td>4</td>
<td>1.58</td>
</tr>
<tr>
<td>Categories X Age X Order</td>
<td>1.77</td>
<td>12</td>
<td>0.68</td>
</tr>
<tr>
<td>Error(^a)</td>
<td>62.15</td>
<td>288</td>
<td></td>
</tr>
<tr>
<td>Emotions X Categories</td>
<td>2.15</td>
<td>12</td>
<td>1.69</td>
</tr>
<tr>
<td>Emotions X Categories X Age</td>
<td>5.35</td>
<td>36</td>
<td>1.41</td>
</tr>
<tr>
<td>Emotions X Categories X Order</td>
<td>0.80</td>
<td>12</td>
<td>0.82</td>
</tr>
<tr>
<td>Emotions X Categories X Age X Order</td>
<td>4.18</td>
<td>36</td>
<td>1.10</td>
</tr>
<tr>
<td>Error(^a)</td>
<td>91.26</td>
<td>864</td>
<td></td>
</tr>
</tbody>
</table>

\(^a\)Each error term applies to the effects in the group preceding it.

\(^b\)\(p < .001\)

\(^c\)\(p < .01\)

\(^d\)\(p < .05\)
Figure 1. Scores for each category summed across emotions and subjects for each age group.
Scalogram Analysis

As noted in the Review of the Literature, data as presented in Tables 1 through 5 or as evaluated by analysis of variance can suggest a developmental sequence but does not indicate to what extent the same sequence is followed by all individuals. Scalogram analyses were performed therefore to test the hypothesis of an invariant sequence, for each emotion separately and on the overall scores for the total sample and separately by age groups. The results of these analyses are presented in Tables 7 and 8.

Reproducibility (Rep) measures the degree to which a scaling pattern is present, and .90 is generally considered the minimum criterion for the existence of a scale (Guttman, 1950). Green's (1956) Index of Consistency (I) measures the degree to which the scalability obtained is an improvement over the minimum scalability possible, given the proportion of subjects passing each item. Generally, a set of items can be considered scalable if I is greater than .50. Finally, Proctor (1970) has devised a chi-square goodness-of-fit test, which, if significant, allows us to reject a scalogram model at the indicated probability level. Generally, scalogram is found to be an inadequate model by this test if the non-scale patterns are not random.

According to these criteria, as shown in Table 7, only the responses for the emotion happy are found to be
Table 7
Scalogram Analysis by Emotions

<table>
<thead>
<tr>
<th>Emotion</th>
<th>Rep</th>
<th>I</th>
<th>Chi-square Significance Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Happy</td>
<td>.91</td>
<td>.73</td>
<td>NS</td>
</tr>
<tr>
<td>Sad</td>
<td>.85</td>
<td>.59</td>
<td>NS</td>
</tr>
<tr>
<td>Mad</td>
<td>.87</td>
<td>.64</td>
<td>NS</td>
</tr>
<tr>
<td>Scared</td>
<td>.85</td>
<td>.61</td>
<td>&lt;.01</td>
</tr>
</tbody>
</table>

Note: A set of items can be considered scalable if Rep > .90, I > .50, and the chi-square test is not significant (NS).

Table 8
Scalogram Analysis for Overall Scores

<table>
<thead>
<tr>
<th>Age Group</th>
<th>Rep</th>
<th>I</th>
<th>Chi-square Significance Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>5-6</td>
<td>.96</td>
<td>.76</td>
<td>NS</td>
</tr>
<tr>
<td>7-8</td>
<td>.90</td>
<td>.66</td>
<td>NS</td>
</tr>
<tr>
<td>9-10</td>
<td>.96</td>
<td>.87</td>
<td>NS</td>
</tr>
<tr>
<td>11-12</td>
<td>.82</td>
<td>.22</td>
<td>NS</td>
</tr>
<tr>
<td>All S's</td>
<td>.91</td>
<td>.73</td>
<td>NS</td>
</tr>
</tbody>
</table>

Note: A set of items can be considered scalable if Rep > .90, I > .50, and the chi-square test is not significant (NS).
scalable. The responses for the other three emotions approach, but do not reach, scalability in Guttman's sense. In addition, a scalogram model is rejected in the case of scared because of the chi-square test. As Table 8 illustrates, an acceptable scalogram pattern is shown for the overall measure when all subjects are considered together. When subjects are considered separately by age group, only the 11-12-year-old does not show a scalogram pattern. Thus, the overall measure appears to be fulfilling its purpose of providing a score which follows the general pattern of the individual emotion scores but is to some extent an improvement over those scores, possibly because it makes allowances for some measurement error.

Despite these generally positive findings on the scalability of the overall scores, some caution is indicated. To begin with, only 61 out of 80 subjects actually show one of the true scale patterns, leaving approximately one fourth of the subjects showing a non-scale pattern. Furthermore, contrary to one of Guttman's (1950) guidelines, there appears to be some clustering among the non-scale patterns. For example, five subjects passed category B and no other category. In developmental terms, this may mean that a significant minority of individuals conceptualize sequential conflicting toward different objects before they do so toward the same object. The failure of Proctor's (1970) goodness-of-fit test to reject the
scalogram model suggests that these irregularities may be due to random error. But just as we cannot accept the null hypothesis, we cannot truly verify the scalogram model. The safest conclusion would be that most individuals develop ACCE in the same sequence of stages, but some individuals may show a different sequence.

In the previous paragraph, it was suggested that a significant number of subjects attained category B before category A. Earlier it was pointed out, based on the data in Table 1, that categories C and D appeared to be of approximately equal difficulty. Such interitem comparisons are best made using Loevinger's (1947) interitem homogeneity coefficient ($H_{ii}$), which measures the extent to which passing the harder item implies passage of the easier item as well. In this way, it can be determined if some parts of the sequence are more scalable than others.

Interitem homogeneity coefficients for all pairs of categories (using overall scores) are shown in Table 9. Unfortunately, no generally agreed upon criterion for an acceptable $H_{ii}$ exists. Kofsky (1966), in a developmental study similar in format to this one, used .50 as a criterion. Using Kofsky's criterion, we find that all item pairs show a significant degree of scalability. It can also be noted, not surprisingly, that the closer two categories are in the developmental sequence, the lower their interitem homogeneity, whereas perfect homogeneity is
attained by items widely separated in the sequence.

Table 9
Interitem Homogeneity Coefficients for Pairs of Response Categories

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1.00</td>
<td>.90</td>
<td>.74</td>
<td>.66</td>
</tr>
<tr>
<td>B</td>
<td>.87</td>
<td>.69</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>.54</td>
<td>.69</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>.70</td>
<td>.69</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

It was possible to construct a sequence of levels showing almost perfect scalability by examining the inter-item homogeneities as well as the overall scalogram patterns. Level 0 was defined as passing no categories. Level 1 was defined as passing A or B or both, but none of the other categories. Level 2 was defined as passing any of the other categories as well as A and/or B. Only one subject passed another category (E) without having passed A or B, and thus did not fit this pattern. He was assigned to level 2, nevertheless. Thus, although there may not be total invariance in a six-level sequence of ACCE involving the temporality and object dimensions, the three-level model does show almost perfect invariance of sequence.
ACCE as an Overall Structure

In examining the content vs. structure question as it applies to ACCE, we can regard the separate emotions as the content element and the five categories of response, or the three levels formed by combining these categories, as the structural element. If the synchrony model of structural development is applicable, then each response category would represent a cognitive structure such that as soon as an individual could conceptualize a conflict of a particular form between any two emotions, he or she would be able to do so for all emotions. In this case, we would expect to see only all-or-none scores for each response category. That is, either a subject would pass the category for all four emotions or for no emotions. On the other hand, if there is no structural element in the development of ACCE and development for each emotion proceeds independently, then the proportions of all-or-none and mixed (i.e. some emotions passed, others failed) scores for each response category would be predictable from the proportion of subjects passing and failing each separate emotion. Thus, analyzing the number of all-or-none and mixed scores for each response category gives one indication of the structured wholeness of the responses.

Table 10 summarizes the results of such an analysis. Although in no category do we find only all-or-none responses, in each category all-or-none scores do outnumber
mixed scores. Furthermore, the chi-square test of the independence of the emotions is highly significant for each response category. These results, then, suggest neither perfect synchrony of responses for each emotion nor independence of responses for each emotion.

Table 10
Independence of Responses for Each Emotion Analyzed Using Frequencies of All-or-none & Mixed Scores

<table>
<thead>
<tr>
<th>Response Category</th>
<th>Actual Frequencies</th>
<th>Expected Frequencies</th>
<th>Chi-square&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All-or-none</td>
<td>Mixed</td>
<td>All-or-none</td>
</tr>
<tr>
<td>A</td>
<td>51</td>
<td>29</td>
<td>33</td>
</tr>
<tr>
<td>B</td>
<td>49</td>
<td>31</td>
<td>18</td>
</tr>
<tr>
<td>C</td>
<td>46</td>
<td>32</td>
<td>11</td>
</tr>
<tr>
<td>D</td>
<td>43</td>
<td>37</td>
<td>10</td>
</tr>
<tr>
<td>E</td>
<td>45</td>
<td>35</td>
<td>10</td>
</tr>
</tbody>
</table>

<sup>a</sup>For all chi-squares, df = 1, p < .001.

Another way of analyzing the structured wholeness vs. independence of development for the emotions is to compare levels of development for pairs of emotions. Each subject can be assigned to a level of development for each emotion analogous to the levels for overall scores. That is, for each emotion, level 0 was defined as passing no response categories, level 1 as passing A and/or B but no others, and level 2 as passing at least one of the categories C, D, or E.
Tables 11 through 16 illustrate the degree to which subjects show the same level of ACCE for each pair of emotions. While there is clearly not perfect synchrony, there is a tendency for subjects to cluster along the diagonal representing equivalent levels of development. According to Goodman and Kruskal's (1954) review of measures of association for contingency tables, two such measures would be appropriate here—their own gamma and Stuart's $\tau_c$. Both measures have the advantage of taking into account the ordinal nature of the categories. Stuart's $\tau_c$ has the additional advantage of achieving its maximum value of 1.0 only when all subjects fall on the longest diagonal of the table, which is useful for testing a synchrony model. Gamma will be 1.0 if all subjects fall on any diagonal of the table. In addition, $\tau_c$ will be 1.0 only if the cell frequencies in the diagonal are equal, which is not an important constraint in testing this model. Gamma does not have this constraint.

As each measure seemed to have some advantages and disadvantages, both were considered and are shown in Table 17. It was hypothesized that part of the discrepancy between the values of gamma and $\tau_c$ was due to the differential effect of the large proportion of subjects in level 2 for both emotions in any given pair. As this massing of subjects in one cell of the contingency tables is partly an artifactual effect of the age range of the
### Table 11
Bivariate Frequencies for Happy & Sad (All Subjects)

<table>
<thead>
<tr>
<th>Levels of Sad</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>( A_D^{mo} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Levels of Happy</td>
<td>0</td>
<td>2</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1</td>
<td>8</td>
<td>47</td>
</tr>
</tbody>
</table>

\( A_D^{mo} = .75, .50, .13 \)

### Table 12
Bivariate Frequencies for Happy & Mad (All Subjects)

<table>
<thead>
<tr>
<th>Levels of Mad</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>( A_D^{mo} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Levels of Happy</td>
<td>0</td>
<td>3</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>6</td>
<td>8</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>2</td>
<td>8</td>
<td>46</td>
</tr>
</tbody>
</table>

\( A_D^{mo} = .45, .50, .13 \)

### Table 13
Bivariate Frequencies for Happy & Soared (All Subjects)

<table>
<thead>
<tr>
<th>Levels of Soared</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>( A_D^{mo} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Levels of Happy</td>
<td>0</td>
<td>4</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>12</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>2</td>
<td>8</td>
<td>46</td>
</tr>
</tbody>
</table>

\( A_D^{mo} = .75, .43, .10 \)
Table 14
Bivariate Frequencies for Sad & Mad (All Subjects)

<table>
<thead>
<tr>
<th>Levels of Mad</th>
<th>AD_{mo}</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 1 2</td>
<td></td>
</tr>
<tr>
<td>0 3 1 0</td>
<td>.25</td>
</tr>
<tr>
<td>1 4 9 9</td>
<td>.68</td>
</tr>
<tr>
<td>2 4 7 43</td>
<td>.28</td>
</tr>
</tbody>
</table>

AD_{mo} \ = \ .77 \ .47 \ .17

Table 15
Bivariate Frequencies for Sad & Scared (All Subjects)

<table>
<thead>
<tr>
<th>Levels of Scared</th>
<th>AD_{mo}</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 1 2</td>
<td></td>
</tr>
<tr>
<td>0 3 0 1</td>
<td>.50</td>
</tr>
<tr>
<td>1 3 13 6</td>
<td>.41</td>
</tr>
<tr>
<td>2 2 8 44</td>
<td>.22</td>
</tr>
</tbody>
</table>

AD_{mo} \ = \ .75 \ .38 \ .18

Table 16
Bivariate Frequencies for Mad & Scared (All Subjects)

<table>
<thead>
<tr>
<th>Levels of Mad</th>
<th>AD_{mo}</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 1 2</td>
<td></td>
</tr>
<tr>
<td>0 5 4 2</td>
<td>.73</td>
</tr>
<tr>
<td>1 0 10 7</td>
<td>.41</td>
</tr>
<tr>
<td>2 3 7 42</td>
<td>.25</td>
</tr>
</tbody>
</table>

AD_{mo} \ = \ .75 \ .52 \ .22
Table 17
Measures of Association for Levels of Pairs of Emotions (All Subjects)

<table>
<thead>
<tr>
<th></th>
<th>Happy</th>
<th>Sad</th>
<th>Mad</th>
<th>Scared</th>
</tr>
</thead>
<tbody>
<tr>
<td>Happy</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sad</td>
<td>.38</td>
<td>.83</td>
<td></td>
<td>.79</td>
</tr>
<tr>
<td>Mad</td>
<td></td>
<td></td>
<td></td>
<td>.71</td>
</tr>
<tr>
<td>Scared</td>
<td>.86</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: Tau_c's are shown in upper right and gammas in lower left. For all tau_c's and all gammas, p < .01.

sample, contingency tables were also set up including only subjects in the two lower age groups, so as to increase the balance of the cells. These are shown as Tables 16 through 23, and measures of association for these tables can be found in Table 24.

Several points are evident from the contingency analysis. Visual inspection of the tables, whether for the whole sample or the restricted age range, will show that the largest number of subjects do fall in the cells representing equivalent levels of ACCE for the two emotions. However, there are enough individuals who do not show such equivalent levels that a model of strict structural synchrony does not seem warranted. Statistically, all of the measures of association are significant. The
Table 18
Bivariate Frequencies for Happy & Sad (Ages 5-6 & 7-8)

<table>
<thead>
<tr>
<th>Levels of Happy</th>
<th>Levels of Sad</th>
<th>( \text{AD}_{mo} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 2 3 0</td>
<td></td>
<td>.40</td>
</tr>
<tr>
<td>1 1 10 6</td>
<td></td>
<td>.41</td>
</tr>
<tr>
<td>2 0 3 15</td>
<td></td>
<td>.17</td>
</tr>
<tr>
<td>( \text{AD}_{mo} )</td>
<td>.33 .38 .29</td>
<td></td>
</tr>
</tbody>
</table>

Table 19
Bivariate Frequencies for Happy & Mad (Ages 5-6 & 7-8)

<table>
<thead>
<tr>
<th>Levels of Mad</th>
<th>Levels of Happy</th>
<th>( \text{AD}_{mo} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 3 1 1</td>
<td></td>
<td>.60</td>
</tr>
<tr>
<td>1 6 8 3</td>
<td></td>
<td>.53</td>
</tr>
<tr>
<td>2 1 5 12</td>
<td></td>
<td>.58</td>
</tr>
<tr>
<td>( \text{AD}_{mo} )</td>
<td>.40 .43 .42</td>
<td></td>
</tr>
</tbody>
</table>

Table 20
Bivariate Frequencies for Happy & Scared (Ages 5-6 & 7-8)

<table>
<thead>
<tr>
<th>Levels of Scared</th>
<th>Levels of Happy</th>
<th>( \text{AD}_{mo} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 4 1 0</td>
<td></td>
<td>.20</td>
</tr>
<tr>
<td>1 2 10 5</td>
<td></td>
<td>.41</td>
</tr>
<tr>
<td>2 1 6 11</td>
<td></td>
<td>.44</td>
</tr>
<tr>
<td>( \text{AD}_{mo} )</td>
<td>.57 .41 .31</td>
<td></td>
</tr>
</tbody>
</table>
Table 21
Bivariate Frequencies for Sad & Mad (Ages 5-6 & 7-8)

<table>
<thead>
<tr>
<th>Levels of Mad</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>AD&lt;sub&gt;mo&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Levels of Sad</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>9</td>
<td>3</td>
<td>.44</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>5</td>
<td>13</td>
<td>.52</td>
</tr>
<tr>
<td>AD&lt;sub&gt;mo&lt;/sub&gt;</td>
<td>.60</td>
<td>.36</td>
<td>.19</td>
<td></td>
</tr>
</tbody>
</table>

Table 22
Bivariate Frequencies for Sad & Scared (Ages 5-6 & 7-8)

<table>
<thead>
<tr>
<th>Levels of Scared</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>AD&lt;sub&gt;mo&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Levels of Sad</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>.31</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>6</td>
<td>14</td>
<td>.38</td>
</tr>
<tr>
<td>AD&lt;sub&gt;mo&lt;/sub&gt;</td>
<td>.64</td>
<td>.35</td>
<td>.13</td>
<td></td>
</tr>
</tbody>
</table>

Table 23
Bivariate Frequencies for Mad & Scared (Ages 5-6 & 7-8)

<table>
<thead>
<tr>
<th>Levels of Scared</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>AD&lt;sub&gt;mo&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Levels of Mad</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>5</td>
<td>4</td>
<td>1</td>
<td>.60</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>9</td>
<td>5</td>
<td>.36</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>4</td>
<td>10</td>
<td>.50</td>
</tr>
<tr>
<td>AD&lt;sub&gt;mo&lt;/sub&gt;</td>
<td>.57</td>
<td>.47</td>
<td>.44</td>
<td></td>
</tr>
</tbody>
</table>
Table 24
Measures of Association For Levels of Pairs of Emotions
(Ages 5-6 & 7-8)

<table>
<thead>
<tr>
<th></th>
<th>Happy</th>
<th>Sad</th>
<th>Mad</th>
<th>Scared</th>
</tr>
</thead>
<tbody>
<tr>
<td>Happy</td>
<td>.52</td>
<td>.46</td>
<td>.45</td>
<td></td>
</tr>
<tr>
<td>Sad</td>
<td>.87</td>
<td>.43</td>
<td>.54</td>
<td></td>
</tr>
<tr>
<td>Mad</td>
<td>.70</td>
<td>.69</td>
<td>.41</td>
<td></td>
</tr>
<tr>
<td>Scared</td>
<td>.71</td>
<td>.85</td>
<td>.61</td>
<td></td>
</tr>
</tbody>
</table>

Note: Tau_c's are shown in upper right and gammas in lower left. For all tau_c's and gammas, p < .01.

The fact that tau_c continues to be lower than gamma, even for the restricted age range, is an indication, at least in part, of the significant number of individuals who do not fall on the longest diagonal of the tables.

It is worth noting that there are no dramatic differences in the levels of association among the six pairs of emotions. However, there are differences in patterns, which a closer look at the contingency tables reveals. For example, of those subjects not at equivalent levels for the emotions happy and sad, an equal number are at a higher level for each emotion. On the other hand, there appears to be more of a decalage between development regarding happy and mad, insofar as 16 subjects show a higher level for happy and only 7 show a higher level for
mad. In general, where some degree of decalage between two emotions is apparent in the contingency tables, the order is consistent with the ordering of the emotions suggested initially by the proportions of subjects passing each item, as summarized in Tables 2 through 5.

Given that the category scores or levels for the different emotions are not independent of each other but that there is not strict synchrony of development either, it is of some interest whether or not there are periods of consolidation characterized by less variability among the levels on the separate emotions. Although conformity to this kind of model is difficult to assess with cross-sectional data, some useful information can be obtained.

One version of a consolidation model would be that although an individual might attain a given level (say level 1) of each emotion at different times, he or she would attain level 1 on all emotions before going on to level 2 on any emotion. Such a model is not viable for these data inasmuch as 9 out of 80 subjects, or about 11%, are at level 0 on at least one emotion and at level 2 on at least one other emotion. In fact, this is probably an underestimate of the proportion of subjects who at some point achieve level 2 on one emotion before having attained level 1 on all emotions. There is an unknown number of subjects who are now at level 1 or 2 on all emotions who
may have shown a two-level discrepancy in the past.

It is interesting to note that there are only 5 subjects who show a pattern of being at level 1 on all four emotions, suggesting that this cognitive level is not a stable equilibrium position. On the other hand, there are 36 subjects who show a pattern of being at level 2 on all emotions. Unfortunately, we have defined no levels above level 2, so that it is impossible to say whether level 2 does represent an equilibrium position or whether the prevalence of this pattern is an artifact.

The possibility of a convergence at level 2 is also suggested by data on the degree of discrepancy between emotions at each overall level, as shown in Tables 25 and 26. At overall level 0, the modal maximum discrepancy between emotions is 2 (at least one emotion at level 0 and one at level 2). At overall level 1, the modal maximum discrepancy is 1 (individual emotions mostly at levels 1 and 2). And at level 2, the modal maximum discrepancy is 0 (all emotions at level 2). As Table 26 shows, this same pattern is obtained even if only age groups 5-6 and 7-8 are considered.

Another way of assessing changing patterns of variability among structurally related cognitive attainments was suggested by Wohlwill (1973). This involves, first of all, computation of the mean absolute deviation from the modal level (\(AD_{mo}\)) of each attainment for each level of the other
Table 25
Percent of Subjects at Each Overall Level of ACCE Showing Different Maximum Discrepancies Between Levels for Any Two Emotions (All Subjects)

<table>
<thead>
<tr>
<th>Overall Level</th>
<th>Maximum Discrepancy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>27</td>
</tr>
<tr>
<td>1</td>
<td>23</td>
</tr>
<tr>
<td>2</td>
<td>81</td>
</tr>
</tbody>
</table>

Table 26
Percent of Subjects at Each Overall Level of ACCE Showing Different Maximum Discrepancies Between Levels for Any Two Emotions (Ages 5-6 & 7-8)

<table>
<thead>
<tr>
<th>Overall Level</th>
<th>Maximum Discrepancy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>30</td>
</tr>
<tr>
<td>1</td>
<td>26</td>
</tr>
<tr>
<td>2</td>
<td>73</td>
</tr>
</tbody>
</table>
attainment. For example, in Table 11 at level 0 for sad, the modal level of happy is also 0. By averaging the deviations from this modal level, we obtain an $AD_{mo}$ of .75 for happy at level 0 of sad. Note that the $AD_{mo}$'s for happy tend to decrease at successive levels of sad, suggesting a pattern of convergence. Similar patterns are evident for all relationships between pairs of emotions in Tables 11 through 16. However, when only the 5-6 and 7-8 age groups are considered (Tables 18 through 23), as a way of avoiding the artificial massing of subjects in one cell of the contingency tables, this pattern is no longer consistent.

**ACCE and Piagetian Measures of Concrete Operations**

As noted previously, Harter (1977, 1979) hypothesized a positive relationship between the ability to conceptualize multiple emotions and the ability to perform concrete operational tasks, such as conservation and multiple classification. One prediction of this hypothesis in the most general sense would be positive correlations between ACCE and conservation and between ACCE and multiple classification. Table 27 shows product moment correlations for total scores on these measures at each age and for the entire sample.

Although there is a moderate overall correlation between ACCE and conservation, this correlation does not
Table 27  
Correlations Between ACCE and Piagetian Tasks

<table>
<thead>
<tr>
<th>Age Group</th>
<th>Conservation</th>
<th>Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>5-6</td>
<td>.00</td>
<td>.17</td>
</tr>
<tr>
<td>7-8</td>
<td>.55&lt;sup&gt;b&lt;/sup&gt;</td>
<td>.30</td>
</tr>
<tr>
<td>9-10</td>
<td>.29</td>
<td>.36</td>
</tr>
<tr>
<td>11-12</td>
<td>.50&lt;sup&gt;b&lt;/sup&gt;</td>
<td>.27</td>
</tr>
<tr>
<td>All S's</td>
<td>.60&lt;sup&gt;a&lt;/sup&gt;</td>
<td>.52&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

<sup>a</sup><sub><i>p</i> < .01</sub>  
<sup>b</sup><sub><i>p</i> < .05</sub>

appear in all age groups. Whether these age differences are the result of developmental changes or are the result of factors like a restricted range at some ages is difficult to say. Nevertheless, the existence of some significant correlations in a rather narrow age range suggests that the correlation between ACCE and conservation is probably not due only to increases in both variables with age. The correlation of ACCE with classification skills, on the other hand, appears only for the overall sample and not for any of the separate age groups. This suggests that this correlation may be dependent on their mutual relationship with age.

Another way of examining these relationships is through the use of contingency tables. Before this was done, it was helpful to establish the sequential patterning
of the conservation and classification tasks, respectively. Conservation of substance, of weight, and of volume were found to form a strong scalogram pattern in that order (Rep = .97, I = .89). Subjects were therefore assigned to levels of conservation according to the most difficult form of conservation attained, as follows: level 3—volume, level 2—weight, level 1—substance, level 0—none. Although there seemed to be no a priori basis for predicting how these levels of conservation would correspond to levels of ACCE, it was thought that an examination of the resulting contingency table might suggest some relationships. This contingency table is shown as Table 28.

Table 28
Bivariate Frequencies for ACCE & Conservation

<table>
<thead>
<tr>
<th>Levels of Conservation</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>AD_mol</th>
</tr>
</thead>
<tbody>
<tr>
<td>Levels of ACCE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>7</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>.55</td>
</tr>
<tr>
<td>1</td>
<td>10</td>
<td>8</td>
<td>6</td>
<td>2</td>
<td>1.00</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>12</td>
<td>17</td>
<td>11</td>
<td>.67</td>
</tr>
</tbody>
</table>

| AD_mol    | .50 | .54 | .40 | .15 |

Some degree of association is apparent from the table and is supported by the statistics (gamma = .63, p < .01; tau_c = .43, p < .01). Although no particular levels of ACCE seem to be exclusively related to particular levels of
conservation, some qualitative relationships are suggested. One interesting point is that a substantial number of children at level 1 of ACCE show no conservation at all, but almost all of those at level 2 of ACCE show some degree of conservation. There is no indication that the relationship between these two variables involves particular levels of one ability invariably preceding particular levels of the other (and not vice versa) as would be predicted by Selman's (1976b) necessary-but-not-sufficient model. The deviation scores ($AD_{mo}$) suggest a progressive consolidation of levels of ACCE at successive levels of conservation. The conservation scores show no clear consolidation, but do show maximum variability for those subjects at level 1 of ACCE.

Unlike the three conservation tasks, the four classification tasks did not prove to be scalable ($Rep = .86$, $I = .54$, chi-square indicating lack of fit to model at $p < .05$). The apparent order of these tasks from easiest to most difficult was matrices, Some-All, dichotomies, and class inclusion, but this order cannot be considered an invariant developmental sequence. Furthermore, to the extent that there is a degree of invariance in this sequence, it is mainly due to the fact that class inclusion tends to be passed only after the other three tasks are passed. Thus the contingency relationships between classification skills and ACCE were examined separately for each
### Table 29
Bivariate Frequencies for ACCE & Each Classification Task

<table>
<thead>
<tr>
<th>Levels of ACCE</th>
<th>0</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Levels of Matrices</td>
<td>0</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>7</td>
<td>18</td>
</tr>
<tr>
<td>Levels of Some-All</td>
<td>0</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>20</td>
</tr>
<tr>
<td>Levels of Dichotomies</td>
<td>0</td>
<td>9</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>18</td>
</tr>
<tr>
<td>Levels of Class Inclusion</td>
<td>0</td>
<td>8</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>3</td>
<td>6</td>
</tr>
</tbody>
</table>

Note: For each classification task, level 0 indicates a failure and level 1 indicates a pass.

### Table 30
Measures of Association Between ACCE & Classification Tasks

<table>
<thead>
<tr>
<th>Measure</th>
<th>Matrices</th>
<th>Some-All</th>
<th>Dichotomies</th>
<th>Class Inclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gamma</td>
<td>.49&lt;sup&gt;a&lt;/sup&gt;</td>
<td>.52&lt;sup&gt;a&lt;/sup&gt;</td>
<td>.62&lt;sup&gt;a&lt;/sup&gt;</td>
<td>.57&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Tau&lt;sub&gt;c&lt;/sub&gt;</td>
<td>.22&lt;sup&gt;b&lt;/sup&gt;</td>
<td>.24&lt;sup&gt;b&lt;/sup&gt;</td>
<td>.36&lt;sup&gt;a&lt;/sup&gt;</td>
<td>.35&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

<sup>a</sup><sub>p < .01</sub>

<sup>b</sup><sub>p < .05</sub>
classification task. These contingency tables are shown in Table 29, and the summary measures of association for these relationships are shown in Table 30.

The general finding is a low to moderate, but statistically significant, positive association between multiple classifications skills and ACCE. This association is not so specific as to suggest that an individual at a particular level of ACCE will inevitably pass or fail a particular classification task. Most subjects at level 0 of ACCE do tend to fail each of the classification skills, while most subjects at level 2 of ACCE tend to pass them. At level 1 of ACCE, there is a tendency to pass matrices, Some-All, and dichotomies and to fail class inclusion. Relationships in the other direction are harder to specify, but there is no support for the notion that the attainment of certain classification skills is necessary but not sufficient for the attainment of certain levels of ACCE.
DISCUSSION

The sequence of levels of ACCE found in this study is very similar to the sequence found by Harter (1979) for multiple emotions in general. Level 0 represents no ability to conceptualize two conflicting emotions together, level 1 represents the ability to conceptualize conflicting emotions sequentially, and level 2 represents, for the most part, the ability to conceptualize conflicting emotions simultaneously. The main qualification to this schematization is the inclusion of category C, the ability to conceptualize conflicting emotions sequentially toward the same object with only an internal change, with the level 2 tasks rather than the level 1 tasks. In fact, contrary to what was predicted, response category C came at the end of the developmental sequence among the tasks used in this study of ACCE.

To some extent, the precise placement of category C in the scalogram patterns may be due to the fact that it was the most difficult to explain to the children without the use of an example. Thus, there may have been some children who could have conceptualized conflicting emotions in this way, but did not understand the task. However,
while this may have changed the scalogram patterns somewhat, it seems likely that, in any case, category C would have remained closer to the simultaneous categories than the sequential categories.

One possible explanation for this finding is that the internal-external dimension is actually more fundamental to the developmental sequence than is the sequential-simultaneous dimension. This is quite plausible if we look at the way each category of ACCE was commonly dealt with. Category A, by definition, involved an external event to change the subject's feelings. Category B, sequential feelings toward different objects, was also generally dealt with in terms of an external event occurring to cause a change in feelings.

Category E, simultaneous conflicting feelings toward the same object, which comes next in the sequence, actually had both common internal and external solutions. An example of an internal solution would be feeling sad about grandfather's death, but feeling happy at the same time because he was out of his misery. This is considered an internal solution because it involves two different ways of thinking about the situation with no separate event attributable to each feeling.

An example of an external solution to category E would be feeling mad at a friend because of a fight, then after making up, feeling good but still feeling somewhat
angry at him. This is considered an expression of simultaneous feelings because there is a point at which the good feelings and the angry feelings coexist (see Appendix). However, it is an external solution because the angry feelings are associated with one event, the fight, and the good feelings with another event, the making up. (It can also be argued, of course, that this example is also less simultaneous, in some sense, than the previous one.)

Finally, category C, which comes last in the developmental sequence, had only internal solutions, by definition. While this internal-external hypothesis rather directly explains the position of category C in the sequence, it does not obviously explain the order of categories D and E, which was the other discrepancy between the hypothesized sequence and the empirical sequence. In fact the principal solutions for category D (simultaneous, different objects) were external solutions. Yet category D followed category E, which had both internal and external solutions. One possible explanation for this is that the internal-external distinction tended to cause more confusion in the case of category D. In order to deal with this category, the child generally had to think of two events that happened (or at least began) at different times, while the two feelings occurred at the same time. In other words, the external elements were often sequential while the internal elements were simultaneous. Because of a
difficulty in separating these factors out, many children attempted to think of situations in which both events, as well as both feelings, happened at the very same time. Although this was sometimes possible, it often led to ludicrous ideas which were rejected by the children. When it became evident that this was the problem, the interviewer would remind the child that the two feelings needed to occur at the same time, but the two events did not. Rarely did this help the child, indicating that the obstacle was usually the child's difficulty in decentering, not a difficulty in understanding the task.

The importance of this internal-external dimension in the development of cognitions regarding feelings has also been noted by others. Cowan (1978) attributed to Inhelder the term "affective realism," which describes the attitude that feelings are always caused by external events. Affective realism is supposedly characteristic of the pre-operational period and tends to persist until the middle of the concrete operational period.

Nevertheless, there is a question as to whether the internal-external dimension is the principal underlying

---

1 According to Cowan's references, Inhelder's ideas were interspersed with Piaget's articles on affect and cognition in the Bulletin Psychologique, 1954, 7, 143-150, 346-361, 522-535, 694-701. These articles have recently been published in book form (Piaget, 1954/1981), but Inhelder's comments were not included in this monograph.
factor in the development of ACCE or whether it is only one more dimension to be considered along with temporality (sequential vs. simultaneous), object, etc. Further research might attempt to deal with the internal-external dimension more systematically, whereas in this study it provided an explicit distinction only within the category of conflicting sequential emotions toward the same object.

Another question for further research concerning the role of the different dimensions in the overall sequence of ACCE is the status of the object dimension. In the present research, the object dimension seemed to play a secondary role to the temporal dimension. (If we take our previous argument seriously, it may actually be the internal-external dimension, or at least a combination of the temporal and internal-external dimensions, that is primary. But we will put that argument aside for now because of the lack of systematic investigation of the internal-external dimension in this study.) Phrased another way, the levels of ACCE were defined in terms of the temporal dimension, while the object dimension resulted in decalages within those levels. Furthermore, these decalages were not entirely consistent among the subjects in that a significant minority attained category B before category A and category D before category E.

It should also be noted that, unlike the results on the temporal dimension, these results tend to contradict
Harter's (1979) hypotheses. Her preliminary data suggested that for sequential responses, different object responses precede same object responses. Although she found same object responses more prevalent than different object responses among all simultaneous conceptualizations, she hypothesized (apparently on an intuitive basis) that for opposite valence conceptualizations, different object responses would precede same object responses. This writer shared Harter's intuition because it seemed that simultaneous, same object responses involved more conflict than simultaneous, different object responses.

Nevertheless, in the present study, for both sequential and simultaneous conceptualizations, most subjects attained same object categories before different object categories. However, there is some question as to whether this is an accurate reflection of development or is due to the interview methodology used in this study. The interview required the subjects not only to be able to comprehend or show competence in dealing with various categories of ACCE, but it also required them to produce examples in each category. Thus, children may have more difficulty formulating examples for conflicting emotions involving different objects than for a single object without necessarily having more difficulty dealing conceptually with one category than with the other. It should be possible to sort this out using methodology which involves
asking children to interpret and explain situations which are presented to them, rather than to produce their own situations. In any event, the results regarding the object dimension are less consistent, from subject to subject within this study and in comparison with Harter's (1979) results, than the results involving the temporal dimension.

Although the levels of ACCE identified in this study are in general agreement with Harter's principal findings, they conflict with the sequence postulated by Selman (1976b). Let us recall again the relevant levels described by Selman for "subjectivity." At stage 1, the child can conceptualize conflicting feelings at the same time, but not toward the same object. At stage 2, the child can conceptualize conflicting feelings toward the same object, but these feelings are relatively isolated from each other, usually in a temporal fashion (i.e. feelings are conceptualized as sequential). At stage 3, a single object can arouse truly (simultaneous) mixed feelings.

Stage 1 generally corresponds to what has been labeled category D in this study, and stage 3 generally corresponds to category E. Stage 2 could correspond to either category A (sequential, same object, external change) or category C (sequential, same object, internal change), or some combination of the two. It is reasonable to assume that Selman's stage 2 would contain mainly category A responses, both because they are much more prevalent than category C
responses and because the example Selman (1976b, pp.162-163) used for stage 2, when he defined his stages, suggested that something had happened to change the child's feelings. If this is the case, then results from the present study indicating that children can invariably handle category A responses before category D responses would contradict Selman's order. If Selman's stage 2 actually contains more of a combination of category A and category C responses, then our main disagreement would be regarding the best way of conceptualizing different levels of ACCE, and our empirical results would be more difficult to compare.

In fact, such lack of comparability does appear to be at least part of the problem. At one point in his paper, Selman (1976b, p.165) used as an example of a stage 2 response, "A gift of a new puppy can make someone happy because that person wants a puppy and sad because it reminds that person of his or her old dog." In this writer's system, that response would be scored as category E (simultaneous, same object) but Selman did not score this at level 3 (which is roughly equivalent to category E) because he saw the two reactions to the dog as relatively separate or isolated. Such disagreements cannot be resolved on the basis of logic, but might eventually be resolved empirically, according to which best captures the sequence and structure of children's thought.
Several interesting questions could be answered if Selman's method, involving the presentation of social situations followed by standard but open-ended questions, was adapted specifically for use on ACCE, rather than the more general issue of interpersonal relations. This is not to suggest that any method is necessarily superior to others, but rather that there needs to be feedback between different methods intending to measure the same construct. As already discussed, we need methods in which situations are presented to the child as well as ones in which the child formulates the situations. In addition, we need feedback between methods in which the children freely define the categories of response and others in which already defined categories are systematically explored.

It would be helpful to use Selman's methods in the area of ACCE first of all to determine what kind of categorization of responses appeared most reasonable. Following such a pilot study, the results of a further study could be analyzed for invariance of sequence using either a longitudinal design or a scalogram analysis with cross-sectional data. It should be recalled here that although in a longitudinal study Selman (1977) found that the overall interpersonal stages of his model did follow an invariant sequence, he has never tested the individual categories or subcategories, such as subjectivity, for invariance of sequence.
If by using Selman's methods we obtained a sequence of ACCE more like his proposed sequence, then we would know that this sequence is highly dependent on the stimuli used to elicit it or on some other factor correlated with the differences in methods (such as the fact that his method tends to deal with one's understanding of others' feelings, whereas Harter's method, used here, deals with one's understanding of one's own emotions). However, this writer predicts that, although using methods like Selman's might not lead to an exact replication of the sequence found in the present study, a sequence conceptualized either as sequential preceding simultaneous or external preceding internal will be evident no matter what method is used.

Another issue which relates to the sequence of levels in the development of ACCE is the question of when a child is able to distinguish notions of sequentiality and simultaneity in a meaningful way. This issue is not directly reflected in the results because it was not foreseen in setting up the interview, and the scoring system avoided the question. In a sense, the interview and scoring system assumed that at the point that children could deal with two feelings, they could differentiate between sequential and simultaneous conceptualizations.

However, a small number of young children gave responses involving two emotions that expressed neither a
sequential nor a simultaneous aspect. An example of such a response would be, "I was mad when someone hit me and happy when I got an ice cream cone." Such a response shows no clear logical or temporal relationship between the two events or the two feelings. In addition, generally children who gave such responses could not establish a connection or meaningfully determine whether the two feelings occurred at different times or at the same time, when they were questioned.

In the present study, responses such as these were scored as sequential, different objects (category B) for lack of a better category (see Appendix, Rule 3 and Example 1). It was felt that these responses did meet the minimum requirements of conceptualizing two opposite valence emotions, and thus should not be left unscored. Scoring such responses in a separate category would not have been feasible with the present methodology because these undifferentiated responses were not found among older subjects, and scalogram analysis requires cumulative rather than disjunctive responses. In any event, these responses were not common, and their scoring should not raise serious questions about the overall results.

However, such responses are important in understanding the development of ACCE, especially among preoperational children. That such children should have difficulty understanding concepts of time is not surprising in light of
Piaget's (1927/1969, Krafft & Piaget, 1925) work in this area. However, Piaget never studied children's concepts of time in a context like our interview procedure. Thus, results indicating that some 5- and 6-year-old children cannot differentiate temporal relationships in recalling and describing events and feelings are interesting in themselves.

For our purposes, these results imply that the development of effective conceptual manipulation of emotions involves not only progressive coordination of conflicting feelings, but also progressive differentiation of time concepts. This is illustrated not only in the early responses just discussed involving two emotions but a complete lack of temporal differentiation, but also in difficulties shown at later ages. For example, some youngsters showed a good ability to coordinate conflicting feelings, but still showed relative difficulty in differentiating time concepts. These children were able to give examples of simultaneous conflicting feelings as well as sequential conflicting feelings, but were often unsure themselves of which was which. Such subjects required extensive questioning by the interviewer in order to categorize their responses. Other children showed a precise differentiation of temporal concepts, but a relatively poor ability to coordinate conflicting feelings. These children gave clear sequential answers when asked for
them, but were unable to think of level 2 answers at all. Thus, these two abilities are at least partly independent, although there is clearly interaction between them and a need for both in order to deal with conflicting feelings in the most flexible way.

The errors made by children who showed some differentiation of temporal concepts, but an incomplete understanding, are interesting and suggest part of what happens in the development of these concepts. Some children tended to confuse the object dimension with the temporal dimension. For example, a child might give a response that she felt sad about moving to a new neighborhood, then she met a friend and felt happy. She might further indicate that those two feelings happened at the same time because they are both about moving. (Of course, she could justify the feelings as simultaneous if she said that she was still partly sad even though she made a friend, but that kind of response is not at issue here.) In a sense, time is marked by events such as moving. Further exploration may reveal similar, perhaps even more global, conceptions of time underlying the responses of younger children to questions about temporal relationships, responses which at first seem random and indicative of no concept of time at all.

Another kind of error reflected a belief that if one feeling came right after the other, expressed in phrases like, "the very next second," then the two feelings could
be considered simultaneous. Similarly, some children, when expressing sequential relationships, felt the need to describe one feeling as coming "a day later" or "a week later," apparently believing that if it came right after, the two feelings might be simultaneous. Future research might investigate children's abilities to differentiate temporal concepts more finely than they were asked to do in this study. One such distinction would be between a situation which immediately engenders simultaneous conflicting emotions and a situation in which one emotion starts first but is still being felt when the conflicting emotion begins. In any event, there is a continuing need to sort out children's understanding of temporal concepts from their ability to conceptualize conflicting emotions and to study the interaction between these abilities.

The consistent theme of this study has been that stages require both sequence and structure. With this in mind, we could interpret the results of this study to mean that the level of sequential conceptualization of conflicting emotions is only a transition phase, whereas the level of simultaneous conceptualization of conflicting emotions is a true stage of development. This would be based largely on the evidence that there is little consistency among emotions at the first level and much more consistency at the second level.
This conceptualization also has an appealing analogy in the development of conservation. As children begin to develop the ability to conserve, they go through a phase in which they can consider the two relevant dimensions (e.g. height and width) in sequence, but not simultaneously. This typically leads to vacillation, with the child deciding first that the higher one has more, then that the wider one has more, etc. Thus this phase lacks the equilibrium of the previous stage, in which the child only considers one dimension, or the following stage, in which both dimensions are considered simultaneously and coordinated. This lack of equilibrium and the resulting inconsistency of responses and relative brevity of this phase are reasons that it is considered a transition period rather than a stage. Similarly, it may be that once children can conceptualize two emotions as part of the same process, they become aware of situations in which emotions vacillate back and forth. It is then only a short step to be able to conceive of simultaneous conflicting emotions.

The fact that it is level 2 of ACCE that is most clearly associated with passage of the conservation tasks is supportive of this position that level 2 represents the consolidation of a stage. The age of attainment of level 2, which is about 9 to 10 years, is also consistent with the general notion that the ability to conceptualize simultaneous conflicting emotions is an attainment of the
Despite this appealing congruence of ideas, the empirical evidence for a stage notion is actually rather inconclusive. The relationships between ACCE and traditional measures of concrete operations are not strong enough to support a classical stage model. Of course, this kind of empirical finding has been a consistent problem for the concept of overall structures. There is also no support in this study for Selman's (1976b) notion that certain logico-mathematical skills are necessary but not sufficient for certain levels of ACCE. However, it must be admitted that this aspect of the study was not designed as carefully as Selman's research, and thus should not be regarded as conclusive.

Another question is whether the apparent consolidation at level 2 of ACCE is due to the fact that this was the highest level defined in this study. Possible further development in ACCE has been described by Selman (1976b) and implied by Fischer (1980). Selman's stage 4 involves the ability to synthesize two opposing feelings into a single feeling. If we apply Fischer's model to ACCE, we would expect a level even beyond that in which two sets of conflicting emotions could be coordinated. One type of situation might involve coordinating one's own conflicting feelings regarding two opposing conditions (e.g. "If I hit my younger sister, I feel good about getting my anger out,
but I also feel guilty. However, if I don't hit her, I feel proud about controlling myself, but my anger gets pent up{	extasciitilde}). Such a thought process might then lead to a resolution of the problem. Another situation at this level might involve coordinating one's own conflicting feelings with the conflicting feelings of another person. In any event, would we see the same pattern of consistency at level 2 of ACCE if further levels of development are included in the study?

One alternative to stage theory is skill theory, as recently set forth in detail (Fischer, 1980). This theory begins by postulating that, although each skill develops according to a structurally defined sequence, there is no reason to expect synchrony among skills. The generic structural levels, which are applicable to all skills, represent analogies between development in different domains rather than actual cognitive links. Fischer does believe that each individual has an optimal level, which no skill of that individual can surpass. As an individual reaches a new optimal level, there will be a significant decrease in the consistency among the levels of various skills, with some skills developing rapidly in response to this new capability while other skills develop more slowly. However, skill theory postulates no necessary periods of clear consolidation of levels across skills.
While thus de-emphasizing relationships between skills involving unrelated content, Fischer emphasizes careful analysis of the skills necessary and the steps involved in mastering particular tasks. Thus, not only was Fischer's theory helpful in suggesting a further level of ACCE, but his theory also provided a backdrop for our earlier discussion of the relationships between the development of time concepts and ACCE. In other words, his theory suggests that we should expect no particular relationships between analogous levels of unrelated skills, but where a task analysis indicates an intersection between different skill domains, there is a close relationship between performance on that task and abilities in the separate skills involved.

Thus, at this point, skill theory seems to provide a more convenient framework than stage theory does for the development of ACCE. Nevertheless, the distinction between skill theory and a modified stage theory, like Selman's, which allows considerable leeway for variations due to content, is mainly a matter of emphasis. It is difficult to say which will be the more useful model in the long run.

Finally, although this study has emphasized theoretical considerations, it seems appropriate to conclude with a discussion of the implications of ACCE, and more generally the cognitive-developmental approach to social and
affective issues, for the field of mental health. Harter's (1977) work on ACCE stemmed from a concern with children experiencing emotional problems, and Selman (1976a, 1976b, Selman & Jaquette, 1977; Selman, Jaquette, & Lavin, 1977), too, has devoted a good part of his study of interpersonal relations concepts to an understanding of childhood psychopathology. Some generalizations from Selman's (1976b, Selman & Jaquette, 1977) work are that emotionally disturbed children tend to lag behind normal peers in interpersonal awareness skills (of which ACCE is presumably a part) and also tend to show less consistency among social cognitive skills and between social and logico-physical cognition.

One question which a finding such as this raises is whether these cognitive lags in the affective and social domains help to cause the emotional problems or whether they are mainly a result or sign of the problems. Intuitively, both positions make sense and probably both processes occur. In the first case, these delays might render a child less able to deal effectively with various forms of stress. Also, as Selman (1976b) pointed out, a child at a lower social cognitive level is less in tune with his or her peers. In the second case, we might expect that once children develop emotional difficulties, their cognitive efficiency would be reduced especially in those areas closest to their conflicts.
These hypotheses are potentially testable, although probably longitudinal studies on large populations would be necessary. The first hypothesis would predict that children scoring low in affective and social cognition would be more likely to develop emotional problems at a later date. The second hypothesis would predict a regression in the level of affective or social cognition coincident with the beginning of emotional difficulty. It should be noted that Selman and Jaquette (1977) found that those children already identified as having significant emotional problems showed slower than normal development in interpersonal relations concepts over a two-year period, but they showed no regression. However, that study did not compare the developmental levels of those children before and after the appearance of their emotional difficulties.

Cognitive-developmental studies can also help in defining the problems of individual children. As norms are developed for cognitive-developmental skills in social and affective domains, these norms can be used in formulating descriptive diagnoses of emotionally disturbed children. A profile of a child's cognitive skills is of interest both in comparing a child experiencing problems with normal children and in comparing the child's strengths with his or her weaknesses. This is similar to the common clinical practice of analyzing the "scatter" on an
intelligence test or on a battery of tests, but the use
of structurally defined levels adds a significant dimen­
sion to such analysis. What is envisioned here is more
of a clinical use of cognitive-developmental information
than a psychometric approach. Nevertheless, considerable
cautions is necessary in applying this approach. For
example, any attempt to apply the normative data on ACCE
from this study should, at the very least, take into
account that the subjects in this study were middle to
upper middle class and above average in intelligence.
Also, if it is noted that a child in therapy appears to be
dealing with conflicting emotions on a lower level than
would be expected at his or her age, one needs to check
out whether this is related to the emotionally charged
nature of the content being discussed, or whether this
child would show this cognitive lag even with the relative­
ly uncharged topics generally discussed by children in
this study.

Once a picture of the child's strengths and weaknesses
is obtained, a treatment program can be worked out. A
cognitive-developmental approach suggests the general
guideline that, if developmental levels are assessed and
deficiencies are found, the goal should be to raise the
child to the next level of functioning. With that very
general principle in mind, specific techniques can be
developed. For example, Harter (1977) reported on the use
of a kind of pictograph technique to help a child acknowledge conflicting feelings by making the conceptualization of those feelings more concrete.

Cowan (1978) suggested that this same general principle may be applicable to treating schizophrenic children, for example by playing games like peek-a-boo with children who are just beginning to develop the concept of a permanent object, no matter what their age. Selman and Jaquette (1977) discussed how an understanding of social cognitive-developmental levels helped clinicians to conceptualize the difficulties of children at a special school for the learning disabled and emotionally disturbed. This understanding was then used to arrange programs in a variety of milieus (class meetings, activity groups, one-to-one crisis interviewing) so as to encourage growth in interpersonal awareness. It should be noted that the cognitive-developmentalists argue for their approach as an additional tool in understanding, diagnosing, and treating disturbed children, not for the replacement of other clinical conceptualizations and treatments (Cowan, 1978; Harter, 1977; Selman, 1976a).

One more point needs to be made regarding the application of cognitive-developmental concepts to the treatment of emotionally disturbed children. It was apparent in carrying out the interviews for this study that two children could do equally well on this task although one
child would be emotionally involved in the task, while another child would approach it as a purely cognitive task. That is, the second child could speak about feelings without actually experiencing those feelings. Intuitively, it would seem that only the first child, who integrated his or her thoughts and feelings, would gain significantly from this "therapy," although both might receive the same score. This brings us back to the point stated in the Introduction, that this study was concerned with how children think about feelings and not how they feel. Much challenging work thus lies ahead in studying the integration of these two facets of development and how to foster such integration.
SUMMARY

This study was designed as an exploration of the degree to which children's ability to conceptualize conflicting emotions (ACCE) follows a cognitive-developmental stage model. A review of the literature revealed that two of the main premises of cognitive-developmental stage theory are the notions of invariant sequence and overall structure. It was thus hypothesized that: (a) the levels of ACCE would succeed each other in an invariant order; (b) ACCE would form a structured whole in the sense of children tending to be at equivalent levels for four separate emotions (happy, sad, mad, and scared); (c) ACCE would be positively related to Piagetian conservation and multiple classification tasks, as all three kinds of tasks require an underlying ability to consider more than one aspect of a situation at the same time.

The conceptual and methodological definition of ACCE was based on Harter's (1977, 1979) earlier work on multiple emotions. Conflicting emotions were defined as multiple emotions of opposite valence (i.e. one positive emotion and one negative emotion). Harter's interview, which dealt specifically only with the temporal dimension (whether
the emotions were conceptualized as sequential or simultaneous), was revised to focus only on conflicting emotions and to inquire systematically about the object dimension (whether the emotions were directed toward the same object or two different objects) as well as the temporal dimension. For responses that involved sequential conflicting emotions toward the same object, the interview also made a distinction between responses involving a change in the situation (external) and those involving only a change in the subject's thinking (internal). Based on Harter's (1979) results and this writer's intuition, the following sequence was predicted: sequential, same object, external change (A); sequential, different objects (B); sequential, same object, internal change (C); simultaneous, different objects (D); simultaneous, same object (E).

The ACCE interview and the conservation and classification tasks were administered to 80 children, 20 at each of four age groups—5-6, 7-8, 9-10, 11-12. Each age group was divided as evenly as possible into boys and girls. The ACCE scores showed no significant effect for sex or for the order of presentation (sequential questions first or simultaneous questions first). There were significant main effects for response category and age, and a significant category by age interaction. This suggested that the structural response categories of ACCE outlined above form a developmental sequence. However, the empirical sequence
of categories for three out of four emotions and for an overall score was ABEDC, rather than the predicted sequence.

A scalogram analysis revealed that these five categories of response formed an acceptable scale, indicating a strong tendency for this sequence to be invariant among individuals. (This was true for the overall score, but for only one of the four emotions when they were considered separately.) An almost perfect degree of invariance was achieved by collapsing the categories into three levels—no categories passed (level 0), A and/or B passed (level 1), E, D, and/or C passed in addition to A and/or B (level 2). This three-level system closely paralleled Harter's (1979) main finding that children are not able to conceptualize multiple emotions simultaneously unless they can also do so sequentially.

The most significant unexpected result in the obtained sequence was the grouping of category C with the simultaneous responses rather than with the sequential responses. This was discussed in terms of the possibility that the underlying factor in the developmental sequence may be the external vs. internal conceptualization of emotions rather than the temporal dimension per se. However, more research is needed to reach any conclusions regarding this issue.

Several different analyses pointed clearly to the fact that development of ACCE is not strictly synchronous
for the four emotions studied. A main effect for emotion in the ANOVA and contingency table comparisons for pairs of emotions suggested some consistent decalage among the emotions. However, the apparent order of development for the emotions—happy, sad, mad, scared—was the same as the order of presentation of the questions in the interview, suggesting that this may be a methodological artifact.

Despite a lack of perfect synchrony among the emotions, there was evidence that responses for the four emotions were not independent of each other. There was a tendency for children who could conceptualize conflicting emotions in a particular way or at a particular level for one emotion to be able to do the same for other emotions. It also seemed that the consistency among responses for the different emotions was greatest for individuals at overall level 2. Alternative explanations for this finding were that level 2 was an equilibrium position or that this resulted merely from the fact that level 2 was the highest level defined in this study. Further research involving higher levels of ACCE was recommended to clarify this point.

There was a general finding of a positive association between ACCE and Piagetian reasoning skills of conservation and classification, but these correlations did not tend to be very strong. There was no indication that conservation
and classification skills were more basic than ACCE, in the sense that certain levels of those logico-physical skills would be necessary for certain levels of ACCE. Children at level 2 of ACCE generally showed at least some ability to conserve and most of them were able to pass the class inclusion item (the most difficult of the classification skills). The possibility was raised that level 2 of ACCE represented an achievement of the concrete operational stage of development. However, overall structural relationships among tasks with unrelated content did not seem strong enough to support a traditional stage model. Thus, this study was consistent with much of the cognitive-developmental literature in finding more support for the notion of invariant sequence than for the notion of overall structure. These results can be accounted for either by a modified stage theory which takes account of both structure and content (e.g., Selman, 1976b) or by skill theory (Fischer, 1980), which emphasizes separate sequences in different conceptual areas.

Finally, this study was discussed in terms of its implications for the diagnosis and treatment of childhood emotional problems. Further study was recommended on the integration of thoughts and feelings.
APPENDIX

Rules

The following is an attempt to make explicit some of the rules used implicitly in scoring the protocols.

1. Only three kinds of responses do not get scored at all: responses such as, "I don't know," "I can't think of any," "That's not possible"; responses involving only one emotion; responses involving two positive valence emotions or two negative valence emotions.

1A. If a child mentions only one emotion, but by its content refers to a previously described emotion, this is scored. For example, let us say that a child is reminded that one thing that made him happy was going to the circus, and then he is asked for sequential conflicting feelings. His response, "I felt sad when the circus was over," would be scored as category A because this is clearly connected sequentially with the happy feeling. However, the response, "I felt scared when there was a fire," would not be scored unless upon further questioning the child could either repeat the two emotions together ("I felt happy when... and I felt scared when...") or in some way indicates a connection between the circus and the fire.

1B. Although two same valence emotions are not scored, a score can be given when a child talks about feeling good then "worse" or, more commonly, feeling bad then "better" (e.g. "I felt angry when my sister tore up my paper, then I felt a little better when she said she was sorry").

1C. The particular emotion does not need to be mentioned by the child if it can be inferred from the context. For example, if a child is asked how he could feel happy and have a bad feeling at the same time, and he responds, "I was enjoying the football game but it was raining at the same time," this is an acceptable response. The interviewer might ask about the particular emotions for clarity, but this is not necessary for scoring.
2. All other responses are scored in one category, and one category only, unless there are clearly separate statements in the response. If a single response fits more than one category, it is scored in the category that is being asked about. For example, the response, "I felt angry when my sister ripped up my paper. Then I felt kind of better, but still a little angry, when she said she was sorry," could be scored as either sequential or simultaneous, depending on what the question was.

3. Responses which contain two opposite valence emotions, but no apparent relationship, either sequential or simultaneous, are scored as sequential (e.g. "I was happy when I went to the circus and I was scared when my house was on fire").

4. The notion of simultaneity is defined in the loose sense that the two emotions must coexist at some point. Thus, responses which suggest that one feeling began first, but the child was still experiencing that feeling when the second feeling began, can be scored as simultaneous in this sense.

5. While it is necessary to consider the overall context of a response, certain key words usually provide a good clue as to whether the response expresses sequential or simultaneous emotions.

5A. Words usually implying simultaneity are: "also," "still," "then again."

5B. The words, "then" and "and then" usually imply sequence.

5C. The words, "and," "but," and "but then," are ambiguous. The thought in, "I felt happy and scared..." is probably simultaneous, but see Rule 3 for an example with the word "and" scored as sequential. "But then" expresses sequence when "then" means afterwards, but it expresses simultaneity when "then" means in addition.

6. When a response is ambiguous as to category, and the subject is asked to categorize the response, his or her decision is usually accepted unless the explanation of the response contradicts this decision. This is particularly important in determining whether a response refers to two separate objects or to two aspects of the same object. Many responses can objectively be considered in either category, and it is the subject's conceptualization that we are interested in.
7. The category sequential, same object, internal change is scored very strictly. There should be no event whatsoever (i.e. nothing that happens, nothing done to the subject, nothing done by the subject, nothing heard or seen by the subject) in between the first feeling and the beginning of the second feeling. The ending of an event is also considered an event, and is therefore scored as an external change (e.g. "I was happy during my birthday party and then became sad when it ended").

7A. Occasionally a subject might describe a situation in which a feeling began to change due to an internal process, then this change was finalized by an event. This can be scored as an internal change.

8. Generally, it is not difficult to decide which emotion to score a response under. Subjects usually respond appropriately to the emotion being asked about.

8A. However, in the sequential categories, if the subject changes the order of the emotions, the response is scored under the emotion which occurs first.

8B. In the simultaneous categories, it does not matter which emotion is mentioned first. The response is scored under the emotion being asked about as long as it is one of the emotions mentioned.

8C. If, in a simultaneous response, the child mentions two or more of the four emotions at issue in this study, the response is still scored only under the emotion being asked about. However, if he or she alludes back to that response, it can be scored again. For example, the child, asked about simultaneous happy and bad feelings, says that she was both happy and angry when she landed a good punch in a fight. This would be scored under the emotion happy. However, later, when asked about simultaneous mad and good feelings, she says, "That would be the same one I told you about for happy." This same response could now be scored again.

9. Responses are scored only according to their formal properties, and no attention is paid to whether or not the child has ever felt that way or to the likelihood that anyone could ever feel that way. Even responses that are totally in the realm of fantasy are accepted, whether or not the child presents them as fantasy.
Examples

The following examples are taken verbatim from actual protocols to illustrate some of the above scoring rules as well as the style of the interviews. In each case, "I" indicates the interviewer, and "S" the subject.

1. I: "What I'd like you to tell me now is how you can have two different feelings. Let's start again with happy. You told me that one thing that would make you happy is getting a birthday cake, and you said that's a good feeling. Could you tell me how you could be feeling happy and then have a bad feeling after you feel happy?"

S: "A bad feeling?"

I: "Yeah, first tell me about a happy feeling, it could be about a birthday cake or something else, and then tell me how you can have a bad feeling right after that."

S: "When someone spits at you."

I: "What would be the good feeling you'd have first, the happy feeling?"

S: "When someone takes me away to Florida."

I: "Could you tell it to me this way then, so I understand how they happen together. Could you say, 'Someone takes me to Florida, and I feel happy, and then after that someone spits at me, and I feel bad.' Can you repeat that back?"

S: "Yeah."

I: "OK. Tell me what I just said."

S: "Someone takes me to Florida, and I feel happy."

This amount of help, actually providing the model for the child to repeat, would be given only once during the interview, if a child gives only single emotions. In this case, even with this degree of assistance, the child is not able to verbalize two emotions together, and the response is not scored (Rule 1). Although this child experienced similar difficulty for the rest of the sequential questions, when we got to the simultaneous questions, her responses changed.
I: "You told me that one thing that makes you mad is when someone hits you, and that was a bad feeling. Could you tell me how you could have a mad feeling and a good feeling at the very same time?"

S: "When someone spits at you and when someone takes you for a ride."

I: "Do they happen at the very same time or at two different times?"

S: "Two different times."

I: "Which happened first?"

S: "The mad."

This response is scored in category B under the emotion mad. Scoring this as a sequential response can be justified either on the basis of the initial response, which is undifferentiated in regard to time (see Rule 3), or on the basis of her elaboration in response to questioning. The former criterion is probably more reliable in that this same girl, a little later in the interview, gave a similar response, but she claimed those feelings happened at the same time. As the context made this seem unlikely, and she was unable to explain how those feelings could be simultaneous, that response also was scored as sequential. Note that in the response shown above, the fact that the subject does not at first label the mad feeling and never identifies the good feeling is ignored (Rule 10).

2. I: "You said that when your sister does stuff to you, that's one thing that makes you mad. Tell me how you could be feeling mad and then have a good feeling after that."

S: "I could be mad at her, and then someone could call and ask me to come over, and then I wouldn't be mad at her anymore. I'd forgive her. Sometimes she starts crying when I get mad. Like she'll bug me, and I'll get mad at her. Then she'll start crying, and I sort of don't feel mad at her anymore."

I: "Now that was kind of two examples. The first one was about two completely different things. The phone call was what made you feel better. The second time your sister started crying, something happened, and then you felt a little differently about your sister. When your sister started crying, how would you say you'd feel towards her then?"
S: "I'd feel sort of sorry for her."

I: "Is that a good feeling when you feel sorry for her, or is that another bad feeling?"

S: "It's sort of in-between."

This is an example of what is referred to in Rule 2 as two separate statements in a single response, and thus it receives two scores—category B and category A, respectively. Also, in the second scorible response, although the subject does not express a clear positive valence response after the negative valence response, he does appear to be feeling better than before, and thus the response is scored (Rule 1B).

3. I: "Another kind of thing that might happen is where you'd be feeling happy about one thing, and at the very same time, you'd have a bad feeling about something different, so that the two feelings would be about two different situations, but they'd happen at the same time. Could you think of an example for that?"

S: "Hmm, maybe it was... I can't think of one right off the bat."

I: "Well, you could use the example of the birthday for the happy one, or you could think of something else."

S: "Well, I'm not exactly sure."

I: "Do you think that could happen, though, that you could feel happy about one thing, and at the very same time you're feeling happy about that, have a bad feeling about something different?"

S: "Yeah, it could happen. Maybe, I could have a fight with my Mom before my birthday, and then my birthday came, and I started feeling happy again, but I still felt sad."

This response was scored as simultaneous (category D). This fits the "loose" definition of simultaneous described in Rule 4. This response also could have been scored as sequential in that the happy feeling was conceptualized as beginning after the sad feeling. However, the simultaneous implication takes precedence because that was what was being asked for (Rule 2). Note also that if this was scored as sequential, it would have been necessary to score it under the emotion sad. However, as a simultaneous response, it was scored under happy, the emotion asked about.
(Rules 8A and 8B).

4. I: "I'm wondering if you could think of one where you're feeling happy about something, and at the very same time, you have a bad feeling about the same thing that's making you happy."

S: "Like I would, I would, I was playing with my friend a long time ago, and then that day... no not... she was going to move away. So at the same time she told me she was going to move away, and I turned really sad because she was my best friend."

I: "Explain that one a little bit more to me. I was thinking, you 'turned' sad. That sounds like you were all sad; first you were happy, and she told you she was going to move, and you got sad. Could you explain how those two feelings would...?"

S: "I have a different one. I have this friend who lives far away, and she came over to our house to visit. Her mom and my mom are best friends. She came to our house to visit, and we were playing out on the playground, and then her mom said, '(friend's name), come on, it's time to go.' And that time I got sort of sad because they had to leave. Is that...?"

I: "It could be. I can't always tell from the situation. That's why I ask you to explain it a little more. It depends on how you would feel about it. So if the way you would feel would be happy first and then when you got the bad news, you would feel sad, that's more the kind of situation we were talking about before."

S: "What did I say before? Then I would know."

What followed was a clarification of the intent of her question, which was to recall a previous acceptable response for simultaneous feelings. This response was repeated for her, but she still could not produce another response expressing simultaneous conflicting feelings. Both of the responses above were scored as category A because her wording suggested sequential feelings, and she could not explain how they might be simultaneous.

5. I: "You said that one thing that would make you mad is if someone threw a rock in your window, and that would be a bad feeling. Again, using that example or making up a new one, let's see if you can think of a way that you could feel mad and have a good feeling at the very same time that you're mad."
S: "Well, if someone threw a rock in my window, I would be mad that they made a mark through it, but if the rock had a note on it saying that I won $100 or something like that, that would make me feel good."

I: "Let's see, for that one, would you say those two feelings are about the same thing or about two different things?"

S: "Two different things."

I: "What would be the two different things?"

S: "Well, it could be winning $100 and having a rock come through your window."

This was scored as category D (two different objects) in accord with the subject's conceptualization (Rule 6). If she had said both feelings were about the rock, it would have been scored as category E (see the following example). Note also that the unlikelihood of this event is ignored in scoring (Rule 9).

6. I: "Now could you think of a situation where you'd be having a happy feeling about something, and the bad feeling would be about the same thing that you're feeling happy about, at the same time."

S: "You could be, I don't know if this is what you want, but you could be happy, well maybe it's about playing again, but you could be outside playing and it's raining. Your mom allows you to go outside. Maybe you wanted to do something with your friends, but they couldn't come out when it was raining. And you could be happy about playing but sad because it's raining."

I: "Well, you tell me, would that be part of the same..."

S: "Yeah, they'd be the same, but is that what you want?"

I: "In what way would those two feelings be about the same thing? What would you say those feelings are both about?"

S: "They're both about the playing."

This example is the converse of the one preceding it, in that here the subject seems, at first, to be referring to two objects, but he is able to explain it in terms of a single object—scored category E.
7. I: "Let's go on and do some examples for sad, then. And again, the example you gave me before was about someone hitting you over the head, so you can use that or something else. Think of a way that you can feel sad and then have a good feeling after you feel sad."

S: "I could be sad that maybe one of my best toys was ruined, but I could be happy at the same time thinking about how I ruined somebody else's toys and how they forgave me."

I: "Now, you said, 'at the same time.' I was thinking of an example where you'd feel sad first and then you'd have the good feeling afterwards. I'm wondering, could you use the same example again? Could that happen where they happen one after the other instead of at the same time?"

S: "Yeah, like first I feel sad that you broke the toy, but then you'd look so sad and feel so sorry that I'd forgive you and feel happy."

The first response is clearly a simultaneous response. No inquiry was done as to whether she thought of this as one object or two, because we had done the simultaneous questions already, and she had given acceptable answers for both relevant categories. The second response may at first seem to represent an internal change (category C), because she has re-thought the situation. Nevertheless, she is responding to an external stimulus (the reaction of the other person), and this response is scored category A. Again, the probability of these feelings is not judged.

8. I: "Let's give another try at a situation where you're feeling mad, and nothing happens, but you think about it a little differently, and then you have a good feeling."

S: "Yeah, my friends walk off and leave me, you know, just sitting around, and then I just think about it..."

I: "When your friends walk off, how do you feel?"

S: "Sad."

I: "That's what I thought you were going to say. We're talking about feeling mad now. Can you think of one for mad, or do you want to go back and do one for sad first?"

S: "No, I'll do mad. My friends talk about me, and I'd
feel really mad, and I'd just think about it, and then I'd say, 'Well, they're still my friends. I'll go make up with them and be their friend.' So I'd just... I mean I never leave my friends."

I: "Now, let me ask you this. When you were giving me that example, were you thinking that you would feel better just after you thought about it, or would you really not feel better until you went over and made up with your friends?"

S: "Really not feel better until I thought about it and... really sort of both. Like I'd feel a little better if I thought about it, and I'd feel a lot better whenever I made up."

This is scored as an internal change (category C) because the change of feeling was begun before the event (making up) occurred (Rule 7A).
REFERENCES


Piaget, J., & Inhelder, B. *The child's construction of quantities: Conservation and atomism.* London:


Selman, R. L., & Jaquette, D. Stability and oscillation in interpersonal awareness: A clinical-developmental


