LANGUAGE EFFECTS IN
ASSESSMENT OF CLASS INCLUSION ABILITY

by

BARBARA HODKIN, B.A.

A Thesis
Submitted to the School of Graduate Studies
in Partial Fulfilment of the Requirements
for the Degree
Doctor of Philosophy

McMaster University

LANGUAGE EFFECTS IN
ASSESSMENT OF CLASS INCLUSION ABILITY

DOCTOR OF PHILOSOPHY (1978)
(Psychology)

McMASTER UNIVERSITY
Hamilton, Ontario

TITLE: Language Effects in Assessment of Class Inclusion Ability

AUTHOR: Barbara Hodkin, B.A. (University of Colorado)

SUPERVISOR: Dr. Linda S. Siegel

NUMBER OF PAGES: xiii, 186

Abstract

Classification, the recognition, construction, and ordering of classes, is a basic part of thought process. One aspect of classification, class inclusion, requires the quantitative comparison of a superordinate class and its included subclass. According to Piagetian theory, class inclusion is the definitive test of classification competence because inclusion ability demonstrates understanding of the relationships between different hierarchical levels of classes.

Consideration of the relationship of language and thought leads to critical examination of the role of language in assessment of classification skills. To the extent that classification ability precedes or is independent of language development, procedures which rely on linguistic comprehension and production will yield results which are incomplete and possibly invalid. If we vary language in assessment, we should obtain a better understanding of early classification strategies and ability.

This research on class inclusion treats language as an independent variable in two ways, by examining alternate verbal forms of the inclusion question, and by reducing language in that question. Within the Piagetian research tradition there is only one classification study with reduced language, and there is one recent experiment which compared an alternate verbal form to the standard Piagetian question for class inclusion. Several experiments are reported here which examined the effects of language variation on class inclusion performance.

In the first alternate form of the inclusion question the word "all" modified the superordinate class, and children gave more correct answers in this case than with the standard unmodified question. Further examination indicated that children performed better with the "all" question because they understood that question better than the standard question.

With the second alternate inclusion form, Siegel's question, children were asked whether they wanted the superordinate class or the subclass. This was an attempt to replicate a previously reported finding that young children gave more correct answers to this form than to the Piagetian question. In the work reported here, after the children answered they were told to take their choice, and when analysis included both this behavioural measure and verbal responses, there were no differences in performance on the standard versions of Siegel's and Piaget's questions. This discrepancy with the earlier finding is discussed in terms of methodological differences between the studies. In contrast to the standard version results, when the word "all" was used in both questions the performance differences obtained were consistent with the previous finding: younger children gave more correct answers to Siegel's than to Piaget's question.

Language in the inclusion question was reduced by using visual symbols to identify the subclasses and the superordinate class. Four-year-olds gave more correct answers to the reduced language question than they did to the standard question in which words identify the classes.

A mathematical model is presented which estimates underlying components of correct and error response to the inclusion question, allowing analysis of the strategies that children use with different forms of the inclusion question. For example, the four-year-olds did not do better with the reduced language question because they were demonstrating more inclusion logic with that form. The difference in results was based on use of different non-inclusion strategies: with the reduced language question the children were more likely to guess, and with the standard question they were more likely to compare the subclasses to each other. Use of the mathematical model was critical in preventing false interpretation of the results.

The methodological and analytic implications for class inclusion research are discussed. In general, these studies demonstrate the usefulness of alternate approaches to inquiry, analysis, and interpretation of the existing data concerning the development of thought processes.

Acknowledgments

Like Rabbit of Pooh Corner, I am blessed with many friends-and-relations, and I would like to speak of them.

When Linda Siegel agreed to be my adviser, I was pleased to work with a psychologist of her stature. That feeling was justified because her advice about experimental problems, methods, and interpretation has been invaluable. Over the course of time, I came to realize that Linda's assets go quite beyond professional expertise in the usual sense. She is practical and straightforward, and she accomplishes a great deal with a minimum of fuss. She is filled with optimism. She gives generously of herself, and dispenses positive reinforcement as though she is not worried that the supply may run out. Her main strategy for managing students is to believe in them. In other words, Linda Siegel is truly an educator, one who leads others forth. I hope that I can imitate her in some measure.

Many people have enriched my academic experience. Steve Link is a fine teacher, and he has spent many hours sharing his skills and his commitment to sound methodology and analysis in psychological research. Daphne Maurer has a broad knowledge of developmental psychology, including Piagetian theory. She was helpful because of her approachability, and her familiarity with my work. Stanley Heshka has a strong interest in the workings of scientific investigation, and he contributed at every level from operational details to theoretical analysis. Lee Brooks likes to talk about ideas, and he has the annoying and fruitful habit of raising difficult questions.

I enjoyed my undergraduate years at the Denver Center of the University of Colorado. One of my teachers there was Nell Fahrion, who shared her love for psychology and exhorted us to "Hold your theories lightly!" Jim Colwell taught me some important things about teaching, and about the humility which can be afforded by competence.

My comrades and colleagues have given me personal support and intellectual energy. I especially want to mention Jim Alcock, Rosemary Barnes, Maru Barrera, Lila Krishnan, Glen Lawson, Trish Tolkin-Eppel, and Cathy Yarrow. Helen Miller has been a good friend to me, and to my children. Eva Gede helped me to dig in and keep going on this research, and Bob St. Claire-Smith demonstrated the efficacy of behavioural principles by organizing the final assault on the thesis.

I. Ahmad provided statistical consulation, and Michael Bishop drew some of the figures for the thesis. Ginny Hansen typed the first draft, and Bev Pitt typed subsequent drafts of the manuscript. Bev was a good sport about my atrocious work habits, and she is responsible for the attractive appearance of this thesis.

This research could not have been done without the help of the Hamilton-Wentworth Roman Catholic Separate School Board. John Grosso, Professional Development Officer, was efficient and pleasant in expediting contacts with the schools. I appreciate the co-operation of the principals and staff of St. Catherine of Siena, St. Cecelia, St. Columba, St. Joseph's, St. Theresa, and St. Vincent de Paul schools. The directors and staff at Little People's Day Care Centre, St. Mark's Co-op Nursery, and Tapawingo Tribe Day Nursery made it possible for me to work with pre-school children. At each of these schools the

teachers were generous in allowing disruption of classroom activities, and they were often interested in discussing children's learning and development. The parents co-operated by giving permission for their children to take part in the research. Most of all, I thank the beautiful, patient children who participated so willingly in these studies.

The roots of personal motivation are complicated and wide-spread. I remember my grandfather, Frank Snyder, who urged me to "get that piece of paper." My grandmother, Ione Eley Snyder, first introduced me to some important aspects of Western culture, gin rummy and Beethoven, and she proudly showed my university degree to her butcher. My parents, Philip Hodkin and Doris Snyder Hodkin, have encouraged me to find my own way. They both accepted serious disruption of their own lives to help my family during a very difficult period. They moved us half-way across the continent, loaned me money, did household repairs, cared for sick children. They were there. My sisters, Laurel Hodkin and Carol Hodkin Groat, and my brother-in-law, Robert Groat, have helped prepare reports, care for children, sustain spirits, and generally maintain a reasonable perspective on the world.

My daughters, Ann Gushurst, Kristin Gushurst, and Bruna Gushurst, have contributed substantially to this work. Their arrivals in my life focussed my interest on questions of human experience, particularly the developmental aspects of human experience. My children have given me moral support, they have been willing guinea pigs for experimental ideas, and they have joined in discussions about how people think. Their contributions to these discussions were typified by common sense, and unwillingness to accept nonsense or jargon. Bruna never tired of making research materials, and she was diligent in prying me out of bed in the morning. Kristin helped with typing, and she brought her artistic sense to bear on the cutting and pasting of stimulus materials. Ann used her competence to do whatever was needed. She typed, prepared tables and graphs, and assisted in running research.

L'amore mio, Enrico Bisutti, mi porta le rose e ride con me. Rick has an abundance of energy, persistence, confidence, curiosity, enthusiasm, and humor, and he has shared these things with me. This work belongs to both of us.

This thesis is dedicated to Ann, Kristin, and Bruna, whom I love dearly, and to the other children who joined in this research. May you all grow in understanding and joy.

TABLE OF CONTENTS

LIST OF TABLES

Experiment 1:

Experiment 2:

Experiment 3:

LIST OF FIGURES

Experiment 1:

Experiment 2:

Experiment 3:

General Discussion

Chapter 1

Language and Thought

Language and thought are important to us because they seem uniquely human. Some of the other primates solve problems and can be taught to communicate with symbol systems. But only human beings spontaneously develop an extensive language system in their natural social environment, and humans have more complex and abstract thought processes than other animals.

The general purpose of this thesis is epistemological because the focus is on ways of knowing about children's thinking, based on a consideration of the relationship of language and thought. Assessment procedures used in the study of cognitive development are examined with regard to effects of language in our estimation of children's logical ability. Within this general context, the specific inquiry concerns language effects in the assessment of class inclusion ability.

Language is a system of conventionalized signs which communicate feelings and thoughts. In this work language refers to words in grammatical relationships, that is, human verbal language which expresses and symbolizes instances, ideas, classes, and relationships. Thought is difficult to define. Dictionaries speak circuitously of that which is in the mind, the act or process of thinking, judgment and reason. The word "thought" is used here to refer to the cognitive processes of receiving, processing, transforming, storing, and recalling information, and there is special interest in the understanding and manipulation of relations, for example understanding the relationships between classes.

Spoken and written language can be observed and measured in psychological research. Thought is quite a different matter. B.F. Skinner, the well-known behavioural psychologist, does not deal with thought because it cannot be observed. Cognitive and cognitive developmental psychologists do study thought even though it cannot be seen, touched, tasted, or directly measured. They do so by treating thought as a hypothetical construct which is inferred from certain behaviours, including, but not limited to, verbal behaviours. The point of the research reported here is to consider the effects of language in the study of cognitive development, specifically in class inclusion, and to pursue this inquiry we must consider the relationship of language and thought.

The Relationship of Language and Thought

There are three possible relationships between language and thought. One is that language is fundamental, that it leads and informs thought. Another is that thought is prior to language, and that cognitive development is instrumental to acquisition of language. A third possibility is that thought and language are independent. In an extreme form this seems implausible since language is obviously a vehicle for the expression of thought, and less obviously, thought appears to shape and even create language in certain instances, for example in poetry and other aesthetic uses of language. However, it is possible to imagine language and thought as independently functioning systems which are brought to bear upon one another when appropriate. A variation of this position is that language and thought are separate systems, not directly dependent upon each other, but that both are based on some underlying process. In this view language and thought are parallel. In considering the three possibilities, we might bear in mind that the relationship between language and thought may vary depending on the stage of human development.

The Position that Language is Basic to Thought

The position that language is basic to thought was clearly expressed by Benjamin Whorf, who considered speech "the most human of all actions (1959, p.178)." Whorf said, "The background linguistic system...is itself the shaper of ideas, the program and guide for the individual's mental activity, for his analysis of impressions... (1956, p.212)." One implication of this position is that because there are different human languages, thought differs between cultures. Whorf asserted that the grammatical structure of language shapes not only perception, but also the organization and evaluation of experience. Thus, there are different world views and different modes of cognitive functioning, and those differences are language-based. Penn (1972) pointed out that Whorf's position may be interpreted in an extreme form, that language determines thought, or in a milder form, that language influences thought.

Carroll and Casagrande (1958) directly tested the Whorfian assertion that grammatical structure affects cognitive functioning. With some Navaho verbs, for example "to carry", the ending attached to the verb is determined by the physical form of the verb's object, and Navaho children are able to use these verb endings correctly by age 3 or 4. Since attention to form is required in Navaho speech, it was predicted that children who predominantly spoke Navaho would be more likely to use form to group objects than children who predominantly spoke English. Navaho children, aged 3 to 10, were divided into Navaho-dominant and English-dominant groups, and were given a matching task. They could match an object to a model on the basis of either color or form. Consistent with the prediction, the children who were predominantly Navaho-speaking were much more likely than the English speakers to match the objects on the basis of their common form, rather than on the basis of common color. Carroll and Casagrande also tested a control group of white, urban, middle-class children on this task. The white children performed like the Navaho-dominant group: they were more likely to match by form than by color. This additional finding complicates interpretation of the results. Dale (1972) suggested that different factors may predispose different groups to prefer form-matching. Exposure to their language structure may influence Navaho-speaking children to match for form at an early age. For the white children, their higher social-economic status may speed up the transition from color grouping to form grouping, since the usual developmental sequence in sorting is to use first color and then shape.

Brown and Lenneberg (1954) tested the Whorfian position in an experiment based on the fact that color categories are represented differently in various languages. They predicted that people would more readily recognize different shades of a color if those shades have a simple name in their language. For example, yellow-green is named by one word in Shona, but it has a composite name in English. The results supported the prediction that availability of simple verbal labels is related to recognition performance. Colors with a clear-cut name were recognized more readily than other hues. However, in a later study Lenneberg (1961) obtained somewhat discrepant results. He gave university undergraduates a sample of the color blue-green, and supplied one of four verbal labels: blue, green, blue-green, or green-blue. The prediction was that the students who received either of the more commonly-used color names, green and blue, would distort their recall choice in the direction of the verbal label more than the students who received blue-green or green-blue. There was, however, no difference in the recognition performance of the four groups of subjects. Lenneberg concluded that language affects recognition tasks with a strong memory demand, like the tasks in Brown and Lenneberg, but that language is not an important factor in hue discrimination or in relatively simple recognition tasks like those used in the second study.

There have been other studies which examined the Whorfian hypothesis by studying color perception, for example Lantz and Stefflre (1964), and Lantz and Lenneberg (1966). These studies measured verbal productivity, speed of response, and memory for instances (Dale, 1972). Overall, these data do not indicate that language determines categorization of color, and thus they do not support the Whorfian position on language and thought. More recent evidence has in fact indicated that the perception and categorization of color are characterized by cross-cultural regularities despite language differences in naming colors (Heider, 1971; Rosch, 1977). Since people in different cultures categorize colors in basically the same ways but, to some extent, name colors in different ways, these data also do not support the position that thought, or concepts, shape language. The findings indicate that there is a degree of independence between the naming and the categorization of colors.

Two important Russian psychologists, L.S. Vygotsky and A.R. Luria, have ~~taken~~ the position that language is crucial in the development of thought. Vygotsky (1934/1962) held that language and thought ~~are~~ separate in the very young child, that there is a preintellectual stage in speech, and a prelinguistic stage in thought development. Very soon however, language begins to serve cognitive development. The first function of speech is social, but speech then differentiates into a communicative function and an egocentric function. Egocentric speech is transformed into inner speech "which serves both autistic and logical thinking (1934/1962, p.19)." Vygotsky did not consider all thought to be inner speech; rather, thought and language overlap, with each having some separate areas of function. But he considered the contribution of language to early cognitive development substantial: "...the speech structures mastered by the child become the basic structures of his thinking (1934/1962, p.51)."

Building on the work of Vygotsky, Luria and Yodovich (1959/1968) commented on language as a social mechanism that contributes to cognitive growth.

> Language, which incorporates the experience of generations or, more broadly, speaking, of mankind, is included in the process of the child's development from the first months of his life. By naming objects, and so defining their connections and relations, the adult

> creates new forms of reflection of reality in the child, incomparably deeper and more complex than those which he could have formed through individual experience. This whole process of the transmission of knowledge and the formation of concepts, which is the basic way the adult influences the child, constitutes the central process of the child's intellectual development (1959/1968, p.11).

Thus, adult speech to the child is critical in the organization of the child's mental processes. Furthermore, adult speech organizes and regulates the child's behaviour: the child comes to act under the direction of verbal influence and instruction. In time, the child's own speech assumes this regulatory and organizing effect on behaviour. In reporting a speech remediation program with twins who were language retarded, Luria and Yodovich observed that not only the children's speech was improved. "...We observed the beginnings of meaningful play; there arose the possibility of productive, constructive activity...(and) a series of intellectual operations which shortly before this were only in an embryonic state (1959/1968, p.122)." Luria and Yodovich concluded that first the speech of others, and then one's own speech directs behaviour and cognitive functioning.

Roger Brown (1972) also commented on the socializing and cognitive effects of language. While language and thought probably developed together over the history of a cultural group, the situation is different for the individual child. In talking with the child and naming objects in the environment, Brown feels that parents pass on their own cognitive structuring, so that the child is influenced in categorizing his or her experience. Brown supports a weak Whorfian position: to the extent that children are motivated to learn linguistic cultural content, the language of their community assumes a formative role in cognitive growth.

For Jerome Bruner (1956, 1964, 1966), language plays a central role in mature cognition. Young children can categorize objects non-verbally, by putting items together on the basis of common perceptual properties, or by making some kind of common response to items in a category. However there are limitations to cognitive functioning which is tied to perceptual properties and motor responses; this functioning is static because children are manipulating either concrete objects or images of concrete objects. Language provides children with a means of moving from the concrete to the abstract, from particular instances to general attributes. This affords children greater power to act on experience intellectually. "In effect, language provides a means...to represent and systematically transform the regularities of experience with far greater flexibility and power than before (1964, p.4)." Language not only leads to greater flexibility and complexity of cognitive functioning, but in fact, language restructures intellectual operations. Bruner (1966) asserted that language is instrumental in changes from one cognitive stage to another, for example in the change from focussing on perceptual attributes like form and color to functional attributes that define categories like vehicles, furniture and food.

In considering the position that language is basic to thought, the Whorfian hypothesis engendered a great deal of interest and some research, but because of recent research findings this hypothesis is no longer considered tenable (Dale, 1972; Rosch, 1977). The work of

Vygotsky and Luria concerning linguistic control of behaviour continues to generate some interest in North America, for example the research reported by Miller, Shelton and Flavell (1970). Both Brown and Bruner continue to argue for the importance of language in the development and functioning of logical processes.

The Independence Position

John Macnamara (1977) assumed that language and thought are autonomous systems, and asserted that part of a child's business in development is building a representing relationship between the two, that is, mapping language and concepts onto each other. He takes this position becaues he believes the grammatical structure of language is not semantically determined, and is independent of thought, which occurs in the "intentional domain." Macnamara's approach is philosophical, and is based on the work of psycholinguists.

Scholnick and Adams (1973) examined the possible independence of language and thought. They pointed out that if language and thought are related, one might expect acquisition of logical skills before related linguistic competence. (The opposite might occur as well, linguistic competence preceding demonstration of logical skill acquisition. These authors do not discuss this possibility.) If logical skills are not apparent before related linguistic competence (and presumably vice versa), then it would appear that the two kinds of functioning, cognitive and linguistic, develop independently. This was examined with tasks which required various kinds of reversal. Children 5, 6, and 7 were tested for ability to understand passive sentences which reverse the word order of active sentences, ability to repeat backwards the words of active sentences, and ability to reverse the order of cards in a 3 by 3 matrix. Scalogram analysis indicated that there was no particular order of acquisition of the abilities tested: in this study cognitive ability to reverse and linguistic ability to reverse were independent. However, the results obtained may be specific to the tasks which were used.

Siegel (1978) also examined the possibility that language and thought are independent in the young child. She surveyed research which assessed the development of cognitive abilities and the development of language related to the abilities. Some of these studies compared discrimination of one from two objects to use of singular and plural words (Beilin and Kagan, 1969), oddity discrimination to use of the words "same" and "different" (Saravo and Gollin, 1969), choice of a middle-sized item to use of comparative terminology (Koff and Luria, 1973), and demonstration of time concepts to ability to understand "before" and "after" (Weil, 1970). In all of these studies the concepts were demonstrated before the correlative language, and thus Siegel concluded that these concepts are not dependent upon language. Siegel (1976) obtained similar results when she compared children's ability to choose the larger stimulus set in a concept attainment task to their ability to understand the words "big" and "little", and again when she compared the results of verbal instructions to choose the set with the "same number" to results on nonverbal tasks for equality. Substantial proportions of the three- and four-year-olds passed the concept attainment tasks and failed the language test. Less than 3% of the children did the reverse, passing language and failing on concept attainment.

Siegel made two predictions, based on the assumption that either language and thought are independent, or that thought precedes language. The first is that use of language will not facilitate concept acquisition, and the second, that success in solving a problem is not related to ability to state the method of solution. To test the first hypothesis, three- and four-year olds were given concept attainment tasks for magnitude and equivalence, and they were either given verbal cues in the task, or received no verbal cues. For the three-year-olds, use of cues made no difference in performance, but the four-year-olds were more successful when they received verbal cues. On these tasks, Siegel concluded that language and thought were functioning independently for the three-year-olds, but that language comprehension was beginning to serve conceptual functioning for the four-year-olds.

In the same study, Siegel tested the second hypothesis. After the concept attainment tasks, the children were asked to explain how they knew which one to pick. Their answers ranged from "the big one" (correct) to "my mommy told me" (incorrect). Many children were able to choose the correct items in the task without being able to explain why. They could make the choice, but could not produce appropriate language about the choice. Siegel considered this further evidence against the precedence of language over logic in young children, and in fact, these data indicate that in the early years concepts precede language.

The Position that Thought Leads Language

The position that cognitive development is basic to the development of language is held, explicitly or implicitly, by many linguistic, cognitive, and developmental psychologists including Anglin (1977), Clark (1977), Furth (1971), Nelson (1977), Olson (1977), and Rosch and Mervis (1975). Nelson, for example, said that children try to organize the world with both thought and language, so that they can act in the world and make reasonable predictions. Thought comes first, however, and language is a sign of progress in building stability and sense into one's experiences. Nelson assumed that the conceptual system underlies the linguistic system, and "the child's task is to develop conceptions and acquire semantics to match (1977, p.119)." She further pointed out that often what seem to be conceptual errors on the part of the child are in fact either language errors or problems in matching concepts to the language which is made available by adults.

The Genevan school, headed by Jean Piaget, also espouses the view that thought is fundamental to language in early cognitive development. Piaget himself, Barbel Inhelder, and Hermine Sinclair-de-Zwart have specified this position. Because Piagetian theory provides a well-developed, broad and very influential model of children's intellectual growth, some of the basic assumptions and research findings of the Genevan school will be presented here in some detail.

The Piagetian Position

Piagetian theory begins with the position that individuals have inherent potentials for physiological, behavioural, and intellectual functioning. With maturation, which is also innately defined, children are able to interact with the physical and social environment, making possible the realization of their potentials. Readiness to learn is internally controlled and motivated, but action in the environment is necessary for the growth of intelligence. Children use their physical, behavioural, and intellectual structures in the world of things and other people, and through the process of adaptation these structures grow, and new structures appear.

Cognitive operations are mental actions. The operations are initially built upon and closely tied to sensory-motor activities, but in time the operations become autonomous and generalized. That is, for the infant and young child the mental actions involve concrete objects, and the child's actions upon those objects. Over time the operations become internalized, and abstracted from the world of concrete objects. The operations increasingly are decentered, separated from the individual's needs and often misleading perceptions. Eventually operations are co-ordinated into systems with a general and stable nature. Mature operations are internalized, reversible mental actions, characterized by logical necessity. But it is important to remember that the foundation of operations is in physical action on concrete objects, in the realm of experience and perception, which is probabilistic and changeable.

The most important Piagetian operations are conservation, seriation, and classification. Conservation is the ability to understand that quantity is not changed by transformations which are not relevant to the quantitative attribute. For example, the volume of a quantity of water is not changed by pouring it from a tall, thin container into a short, thick container. Seriation is the ability to order items along a quantitative dimension, and Piagetians often test for this operation by asking children to put in order sticks of varying lengths. Classification is the recognition and construction of classes, and for Piagetians classification operations include the ability to understand the relationships between classes; for example hierarchical relationships. Hierarchical classification can be demonstrated by sorting items into subclasses, and then combining the subclasses into superordinate classes.

The cognitive operations require representation, and the representation used is related to the stage of cognitive growth. The various modes of representation include imitation, symbolic play, images, and language. As cognitive operations become less dependent on physical action and concrete objects, there is increasing internal representation of information. Language is a particularly important kind of representation, "but the child's ability to use verbal signs is dependent on the progress of his own thought (Piaget, 1962, p.273)."

Piagetians report that there is a correspondence in the acquisition of language and logic. There are various reports of correlation between logical attainment and corresponding language use. For example, children's ability to understand "some" and "all" is systematically related to their competence at class inclusion, the logical operation requiring manipulation of subclasses and superordinate class. Similarly, children's use of words which refer to increasing size is related to their operational competence at seriation, the logical operation which requires ordering along a quantitative dimension (Inhelder & Piaget, 1964/1969). When children are able to conserve, to understand that a specific attribute of a substance is unchanged by an irrelevant physical transformation, their language is different in predictable ways from the language of children not yet able to conserve (Sinclair-de-Zwart, 1969).

Granted that there is a correspondence between language acquisition and the development of logic, what is the nature of this relationship? Does language lead logic? Does conceptual development lead linguistic accomplishment? Or is it the case that language and logic develop in parallel, with both areas of function dependent on some other underlying process? Piagetians are very clear in asserting that language does not precede or structure logic (Inhelder and Piaget, 1964/1969; Piaget, 1926/1952, 1972/1976). The learning of language in the social environment does not contribute to construction of operations.

Studies of Cognitive Abilities and Related Language

Piagetians offer various kinds of evidence in support of the position that cognitive operations precede accompanying language. For example, when children are able to demonstrate class inclusion ability with a certain kind of stimulus content, they can also spontaneously use language relevant to that stimulus content, but they cannot produce language at a more abstract stimulus level than that of operational competence. Related to this is an example from the area of conservation. Typically children acquire conservation of substance at around seven or eight, conservation of weight at nine or ten, and volume at eleven or twelve. With each new area of conservation attained, children give essentially the <u>same</u> verbal explanation: "It's still the same, because you only made it longer, you didn't add anything, it's longer but thinner..." Since the same language is used in each case, the ability to conserve appears to have no relationship to language, but rather depends only on the qualities of the objects. In another example, if children enumerate out loud the items in sets, they are no more likely to conserve those sets (Piaget, 1972/1976).

<u>Parallel Development of Classification and Seriation</u>

Another argument against the precedence of language concerns the parallel development of the two basic areas of logical operations, classification and seriation. Because the two kinds of operations have "similar turning points," Piagetians conclude "that the development of operational behaviour is an autonomous process (Inhelder & Piaget, 1964/1969, p.290)," and that while language may contribute to the final expression of logical development, language is not central to the growth of operational structure.

<u>Studies with Language Deficient Children</u>

Piagetians reason that if language contributes to the development of logic, one would expect people who are deficient in language to be cognitively deficient. Based on research done with deaf children by Oleron, Vincent, and Affolter, Piaget (1972/1976) and Sinclair-de-Zwart (1969) claim this is not true. Deaf children have the basic logical operations of classification, seriation, and conservation, although they do experience difficulty with conservation of liquids. The sequence of acquisition of cognitive functions is the same for deaf and hearing children, although sometimes the deaf children are slightly behind in the timing of acquisition.

Furth (1971) reviewed 39 studies which compared deaf and hearing children on a variety of cognitive tasks. On 31 tasks no difference was found between the two groups, but on 22 tasks the deaf children did not perform as well as hearing children of the same age. It is difficult to discern any systematic patterns in these results. Furth reported that the deaf children appear to be disadvantaged in discovering symbols, but that they appear to use symbols as well as hearing children, although he noted that one study found poorer symbol use by the deaf. He cited evidence which is consistent with the Piagetian observation that deaf children's problems with conservation are specific to conservation of liquids. Overall, Furth agrees with the Piagetian position. The deaf have the same logical operations as the hearing, and these operations appear in the same order. Delays in acquisition for deaf children are negligible to moderate, and the deaf children often catch up to the hearing at some point even though their language competence is still not equal to that of people who can hear. Furth supported Piaget in asserting that language is not a contributing factor to basic cognitive structure.

Other authors disagree about the amount of developmental delay that deaf children show in cognitive operations. Springer (1976) pointed out that Caouette and Oleron independently found developmental lags of 4 to 6 years with deaf children. Springer reported that the hearing children in his study demonstrated conservation at age 7, the same age reported by the Piagetians, but deaf children did not succeed at the conservation problems until age 10. He concluded that language does play an important role in cognitive development. Thus, there is some inconsistency in research findings concerning deaf children, and in the conclusions drawn.

There are many methodological problems in comparing hearing and deaf children on tests of logical competence. Furth acknowledged the problem of assuming the two groups are the same in every respect except language ability. In discussing the discrepant results obtained in different studies, Best (1973) pointed out that a wide variety of tasks, materials and procedures have been used. Perhaps the most serious methodological weakness in this research is that usually experimenters who can hear are using methods of assessment designed for hearing people, and they adapt these tasks for use with deaf children. Certainly that is the case whenever the Piagetian tasks are used, or adapted. The validity of an assessment instrument is less when it is used for a population other than the one for which it was constructed and standardized (Anastasi, 1976). It cannot be assumed that removing, or lowering, the verbal requirements of a test make that test equivalent for deaf people to the test which is given hearing people. That is a question for psychometric investigation, and a survey of the psychometric literature indicates that this is a very difficult research question which has not yet been resolved.

Like the deaf, retarded children may be relatively more deficient in language than in other intellectual functioning. Furth and Milgram (1965) studied language effects in cognitive assessment with retarded and non-retarded children. The children were given classification questions about seven items, three of which belonged to a class. The items were either presented verbally, as words, or non-verbally, as pictures. The instructions for response were all given verbally, but the response requirements were either non-verbal ("Point to three that go together") or verbal ("How do they go together?"). Thus there were four kinds of questions, defined by presentation and response modes: pictures-pointing, pictures-verbal, words-pointing, words-verbal. The retarded children did less well than the non-retarded on all of the question types that involved language, pictures-verbal, words-pointing, and words-verbal. But on the question that did not use words in presentation or response, pictures-pointing, the retarded children did as well as the non-retarded: correct response was 66% for the non-retarded and 67% for the retarded children. In other words, when the manner of assessing classification was suited to the retarded, they demonstrated classification competence. If only linguistically-based assessment had been used the retarded would have appeared deficient in the classification task. This finding supports the position that language is not necessary to the development of basic cognitive skills, because the retarded were able to classify despite their linguistic deficiency.

The studies with deaf children are relevant to the question of the relationship of language and thought because we assume that deaf children have less language than hearing children. There is a problem, however, because deaf children often have other forms of language than oral and written language, and it is difficult to assess, or control, the "total" amount of language that a deaf child has. This problem was dealt with by Best (1973) who examined class formation ability in four groups of children who varied in relative exposure to communication. The groups in order of access to communication were: hearing children; deaf children exposed to manual language from infancy, receiving both oral and manual (total) education; deaf children not exposed to manual language before school, receiving total education; and deaf children not exposed to manual language, receiving oral education only. The materials used were representations of real life objects like animals, people, and furniture, and geometric figures like cylinders, red squares, and yellow triangles. Best hypothesized that experience with language would make the most difference in sorting real life objects because language is important in transmitting cultural information about these classes, while language is less important in providing information about the geometric figures for which classes are defined by perceptual attributes. The results supported this prediction. With the real life materials, the hearing children performed better than the manual-total group, and the manual-total group did better than the other two deaf groups. However, there was no difference between the groups in performance with the geometric figure material. Best concluded that language deficiency impaired performance with real life categories because language provides important information about those categories. But the results with geometric figure material showed that language deficiency did not impair the development of basic ability to classify.

To sum up these studies with language deficient children, the retarded and the deaf, there are some methodological and interpretive problems with this research, and the data are inconsistent with regard to the amount of delay experienced by language deficient children in the acquisition of cognitive operations. But these studies do indicate that language deficient children acquire the same basic operations as other children, and they acquire the operations in the same order. This suggests that language is not necessary to the development of cognitive operations. Certain results justify a stronger position, that with some operations there is no difference in ability between language deficient and other children, if linguistic difficulties in assessment are removed.

Studies with Verbal Training

Training studies involving verbal procedures are interpreted by Piagetians as offering yet another source of support for the position that language does not structure logical operations. Morf (cited in Piaget, 1972/1976) studied children judged to be in between the concrete and formal stages of operational development. The intellectual tasks are not identified, but involved "some reasoning problems including implications, disjunctions, etc." After the initial assessment, Morf provided the children with verbal expansion of the questions, including more details and similar examples. He found that this verbal procedure improved subsequent problem-solving only for those children who had previously solved at least one problem. The implication is that verbal training was not effective except for children who were already in a state of transition between inability and ability to solve the problems.

In another training study, Sinclair and Inhelder (reported in Sinclair-de-Zwart, 1969) determined whether children were capable of conservation, and then examined language differences between children who could conserve and those who could not. There was no difference in comprehension of language, but there were systematic differences in production of conservation-related language between the children who conserved, and those who did not. These differences included use of comparative or absolute terms, differentiation or lack of differentiation for stimulus dimensions, and co-ordinated description. The experimenters tried to teach nonconservers to use language like the conservers. It was easy to teach dimension differentiation, but the nonconserving children had difficulty learning to use comparative terms, and it was even harder for them to produce co-ordinated descriptions of dimensions. In terms of effect on subsequent ability to conserve, 10% of the nonconservers made progress in operational capability after verbal training. The Piagetian interpretation is that this training was not particularly effective, and the conclusions offered are that language acquisition is not instrumental to operational competence, but rather that mastery of the language related to conservation reflects the development of operational structure.

There is conflicting evidence however. Some non-Piagetian studies have been successful in using verbal training for conservation. Examples are the study of Sigel, Roeper & Hooper (1966) in which children were given verbal training for differentiation of dimensions, multiple classification, and reversibility; and two of the training conditions used by Beilin (1965), in which children were given verbal orientation to dimensions, and verbal rule instruction. In all of these cases, the children who received verbal training showed significantly more improvement in conservation than did the children in control groups.

Summary and Evaluation of the Evidence

The preceding discussion presented various kinds of evidence used by the Piagetians to support the position that cognitive development leads language acquisition. The parallel development of conservation and seriation is a weak argument, because one might posit that the parallel results from language acquisition underlying the development of both conservation and seriation. The Piagetians have reported that verbal training for cognitive operations is not effective, but other researchers have attained success with verbal training studies, and so these results are inconsistent. There are many studies which reported logical competence before the appearance of accompanying language, and no studies which reported the opposite finding of linguistic ability before the accompanying logic is demonstrated. Finally, several experiments with children who are deficient in language found that these children acquire the same logical operations, in the same sequence, as do children with normal use of language. These two last lines of evidence, based on the appearance of logical capability and accompanying language, and on the logical abilities of language deficient children, provide satisfactory support for the Piagetian position that language does not contribute to the structuring of early logical operations, and that language in the young child is a symptom and product of the general cognitive structures[1] (Sinclair-de-Zwart, 1973).

Surveying all of the arguments and data presented concerning the relationship of language and thought, the preponderance of evidence appears to favor the position that early cognitive development either precedes or is independent of linguistic development. At the very least the experimental results justify pursuit of the inquiry into effects of language in assessment of cognitive operations.

Language as an Independent Variable in Cognitive Developmental Research

This discussion concerns methodological issues in cognitive developmental research. Much past research in this area has used methodology that is highly verbal in that language defines the task for the child, and often language is the medium of response. I suspect this is so because adults doing research are themselves likely to use verbal modes of intellectual expression. Asa Hilliard (quoted in Exhibit A, 1977) has pointed out that in the assessment of intellectual functioning we usually ask, "Do you know what I know?" rather than "What is it that you know?" To expand on that, often we ask, "Do you know what I know, and can you show it in the way that I do?" However, as Katherine Nelson says

> ...we only get into trouble when we try to explain development in terms of the child learning the correct (standard, adult) way to do things: to say things, to conceive of things, or to categorize things. This kind of view unduly rigidifies adult modes at the same time

[1]Piagetian theory holds that language may be necessary for completion of the more advanced and abstract operations of the formal level; however, this thesis is concerned with the earlier years, the concrete and pre-operational levels.

> that it denigrates more open and less developed system. but more than that, it actually prevents us from seeing the child's system...it is our job to determine what the nature of the system is, and how it develops... (1977, p.318).

If we want to find out what children know, we must seriously consider the effects of our methodology. Reliance on highly verbal techniques is not consistent with the theoretical position that language is not the source of logic and that language does not precede logic. To the extent that either the independence position or the Piagetian position on language and thought is valid, linguistic aspects of methodology will affect the validity of empirical results in the study of logical development. Use of highly verbal techniques alone may preclude examination of the acquisition of logical operation. A research approach which varies language in assessment procedures should provide the best estimate and the best understanding of developing logical ability. This research strategy of varying language may also provide an additional means of examining the developmental relationship of language and thought.

One way to vary language in assessment is to use different linguistic forms which seem, to the adult, to have equivalent logical meaning. The rationale for this approach is that the meaning of a linguistic form for an adult is not necessarily the same as the meaning for children. If children tend to misunderstand one linguistic form, or understand the form in a different way than adults, then use of other linguistic forms with similar (adult) semantic content may serve to illuminate the nature of the difficulty.

Another way to vary language is to reduce language demands in logical assessment. Language can be reduced by using tasks which are essentially nonverbal. Or, specific aspects of language can be replaced by nonverbal means of communication or representation, either in task instruction or in response requirement.

Miller (1976) has reviewed studies which reduced language in the assessment of Piagetian concepts. These studies reduced language in the presentation of the task and in the response requirements. The more important methodological approaches included motivated choice, in which children point to what they want rather than answering a quantitative question; instrumental choice, in which the child learns, for example, to point to the longer item in a stimulus pair; and use of surprise as a measure of violation of logical expectancies. While finding current results somewhat modest, Miller concluded that the endeavour to reduce language in assessment is worth pursuing. This endeavour may result in more precise evaluation of cognitive development which is of interest in itself, and the results may also bear on questions such as the sequencing of cognitive attainments, stages of cognitive development, and processes which underlie cognitive abilities.

Miller noted that the studies surveyed dealt with the operations of conservation, seriation, and transitivity. There was only one study concerning classification, that of Braine (1962). Braine wanted to examine the acquisition of classification operation, and he presented a well-reasoned argument for the use of nonverbal methods to study this operational development: "It would seem to be intrinsically impossible to study how a concept develops with methods which employ verbal cues to evoke the concept. For, if the child understands the verbal cue, he must already have developed the concept (p.46)." His argument assumes the Piagetian position that concept attainment is prior to associated verbal cues.

To study the development of classification ability, Braine used a similarity-difference discrimination paradigm. Children were rewarded for learning to choose either the stimulus set with similar items, or the other set which contained a variety of items at the next level of hierarchical organization. In one example, the children were to choose either a set of apples (similarity) or the other set with an apple, banana, and peach (difference). Children aged 3 through 6 were able to do the similarity problems quite well (73% correct response) when those problems were of the "species-genus" type like apples and fruit. But the children performed poorly with "genus-family" sets such as fruit compared to various kinds of food. There was also a low rate of success with all of the difference problems.

Braine concluded that the factors that affect children's ability to group items together probably change with age, and that these changes may underlie the development of children's understanding of classification. Miller's (1976) comments indicate that he did not find the Braine results impressive, and, there are criticisms which can be levelled against Braine's procedures and results. I think there are problems with the stimulus items used in terms of representativeness, repetition between sets, and differences in familiarity of the materials between sets. Some of the materials chosen to represent species-genus and genus-family seem questionable. For example, both planes/vehicles and stencils/art supplies were used as species-genus sets; in light of the recent work by Rosch (1976) on natural levels of categories, it is doubtful that planes and stencils are at the same natural category level. More important, Braine's training for similarity and difference discrimination seems to bear little relationship to Piagetian problems for hierarchical classification, in which children are asked to sort, or distinguish, subclasses with very similar members and superordinate classes with exemplars from various subclasses. That kind of hierarchical classification problem seems to be what is intended in Braine's study, but his difference-similarity task is quite different from that problem. In an additional comment on Braine's work, Miller said, "...(the Braine study) suggests, both implicitly in its procedures and explicitly in its discussion, that classification may be one area for which nonverbal techniques are not possible (p.18)." On the second point, he was wrong. Nowhere in the article did Braine deviate from his position that nonverbal methods will be useful in studying the course of logical development, including development of classification. When Miller suggests that there are problems with certain specifics of Braine's procedure it is hard not to agree. However, it is one thing to find that there are weaknesses in the first study attempted in an area, nonverbal assessment of classification ability, and it is quite another thing to conclude that the attempt is futile. I propose that a more reasonable conclusion is that different approaches should be tried to reduce language in the assessment of developing classification competence.

Chapter 2

Language and Classification Operations

The introduction was about the relationship of language and thought, particularly in young children, and the implications that this relationship has for the study of cognitive operations in general. This chapter concerns one area of cognitive operations, classification. The discussion will focus on the nature of classification, and on modes of representation used in categorization, in order to consider the merit of attempting to reduce language in assessment of classification competence. I suggest there are two possibilities regarding the acquisition of classification and language, and thus two hypotheses about reducing language in classification assessment.

Hypothesis I: Language is the medium normally used in the child's own classification operations, and therefore it is inappropriate to try to reduce language in classification assessment. Attempts to reduce language violate the nature of the child's normal classification operation, even in the early stages of acquisition.

Hypothesis II: Early classification ability precedes and is not dependent upon concomitant language forms, and so it is appropriate to try to reduce language in assessment of early classification operation. If language requirements are successfully reduced, a better estimate might be obtained of early stages of classification ability.

Hypothesis I is suggested by the observation that language is, in large part, a classification system. One could further assert that not only is language a classification system, language is the <u>best</u>

example of the generalization and discrimination which constitute classification. One might argue that genuine classification, that which goes beyond recognizing the identity of some specific instances, depends upon language and follows from linguistic attainment.

Roger Brown makes a case like this. He argues that language is a system of "perceptual and cultural categories having important relations with the nonlinguistic categories of thought and culture (1956, p. 312)." Children are informed about concepts from the linguistic utterances and responses of other people, and thus language facilitates the structuring of concepts. With some qualification concerning individual motivation, Brown thinks that language may be formative to cognition in that "the lexical structure of the speech he hears guides the infant in categorizing his environment (1972, p. 255)."

The ideas of Vygotsky and Luria support Hypothesis I. While much of the relevant Russian work concerns the verbal control of behavior, the quotation of Luria and Yodovich in Chapter 1 makes clear that their understanding is much like that of Brown. Adult verbal communication establishes, defines, and structures concepts for children, and this process is basic in cognitive development. This view of language and classification suggests that language is the appropriate vehicle for classification operation, and thus it is not suitable to try to evaluate classification ability with reduced language techniques

Hypothesis II, on the other hand, is supported by a different line of reasoning. As with seriation and conservation, it is assumed that early classification functioning and the language which accompanies classification both depend upon the growth of underlying cognitive structures. In early stages, language use reflects classification development, and early classification ability is in advance of concomitant language. Thus, reduced language assessment will allow better examination of early classification ability.

This position is consistent with Piagetian theory. Inhelder and Piaget (1964/1969) say that "the origins of classification ... can be traced prior to the evolution of language and symbolic representation (p. 15)." Classification structures are mirrored in language, not based in language, and the development of classification operation "is largely independent of language (p. 4)." Rather, classification develops from "sensori-motor schemata", repeated patterns of actions which are used to recognize and to signify objects and situations. It is through acting on the environment and developing action patterns that children attain early classification functioning. The position of the Piagetians, and others who believe that classification ability either precedes or is independent of linguistic attainment, supports the view that it is reasonable to reduce the language used in the assessment of classification ability.

Discussion of Language and Classification[1]

The two hypotheses concerning classification and language reflect the more general discussion of language and thought in the preceding chapter. The hypothesis that language is the normal medium

---

[1]Lee Brooks contributed a great deal to this discussion.

for classification is one part of the position that language is basic to thought. The second hypothesis, that early classification "precedes" and "is not dependent on" language, encompasses the position that thought precedes language and the position that thought and language are independent. I concluded that the evidence surveyed favored these last two positions, but another way of dealing with this issue is to point out that the theoretical positions about thought and language are probably too extreme as stated.

The idea that language precedes thought underestimates the inherent structure of the external world. By default this position implies that all possible combinations of attributes exist in the world, that birds and mammals are equally likely to have feathers and fur, to be red, brown, silver, or green. This view further implies that verbal labels, linguistic categories, are arbitrary. In other words, the "real world" is assumed to be like those studies of concept formation in which nonsense words are assigned to figures with independently varying attributes, for example, squares and circles equally likely to be big or small, to have a dot or a cross in the center, and so on. But this is not the way the world is. In natural objects, attributes occur in correlated clusters; feathers and beaks are likely to be found together, as are fur and claws. Moreover, there are certain focal values among attributes, so that some attributes are more important than others, more likely to be perceived and used in categorization. To the extent that the world is inherently structured, language does not so much define as reflect categories. (For supporting evidence and an extensive discussion of these observations, see Cole and Scribner, 1974, and Rosch, 1977.)

The issue, then, is arbitrariness. The position that language precedes thought assumes extreme arbitrariness in the world, and this leads to the view that children must be taught how to categorize the content of their experience, primarily by means of language. The opposing view asserts that much in the external world is <u>not</u> arbitrary, and so language is not necessary to early categorization. In various disciplines, anthropology, linguistics, psychology, earlier research concerning this issue was influenced by an emphasis on cultural diversity, and this emphasis on differences led to results and interpretations which supported the conceptualization of the natural world as arbitrary. More recently, cross-cultural work has focussed on the similarities between groups of people, and this attention to human commonalities has resulted in awareness of the structure which is inherent in the external environment.

It is not likely that our experience of the world is entirely arbitrary, or that experience is entirely structured in such a way that social learning has no effect on our categorization. The truth surely lies somewhere between these two extreme possibilities, and the problem is determining the degree of structure in the external world. Within the constraints of natural world structure, probably language is a vehicle for socially transmitted information, and a means to shape children's perceptions and categorization by social reinforcement. But to the extent that the world is structured, individual experience will lead to categorization independent of linguistic operations. Because the world is structured to some degree, it is worthwhile to vary language and reduce language in the assessment of classification ability.

Modes of Representation

Since language does not define all classification, some non-linguistic representation must be used for categorization. Piagetians propose that early recognition of object similarity is based on perceptual processes, and classification operations originate with sensori-motor schemata. Jerome Bruner also talks about perceptual and action-based processes, but in a very different way than the Piagetians.

Bruner (1964) describes an evolution in modes of representation. The three systems, in order of appearance are enactive, iconic, and symbolic representation. Each system is based on the preceding one, but all of the systems remain functional throughout life. Enactive representation refers to sequences of sensori-motor actions which represent past learning or experience, and is therefore like Piaget's action schemata. Babies demonstrate enactive behaviour when they shake their fist over and over, after the rattle they have been holding has dropped to the floor. Adults "know" with their fingers how to tie a shoe, but demonstrating this understanding with a drawing, or verbally, is very difficult.

In terms of contribution to cognitive structure, Bruner attributes a higher level of functioning to iconic, perceptually-based, representation than does Piaget. In iconic representation, spatial schemata are constructed from the child's experience with the world.

Information is selectively organized into images which refer to external reality. Examples of this kind of representation occur when we sketch the layout of our backyard or draw a map for a friend, rather than giving verbal information. Bruner considers iconic representation the predominant system for children from 2 to 6.

Around age 6 or 7 children shift to greater use of symbolic representation. As the name suggests, this system represents things with abstract and arbitrary signs. We must be taught that "New York" refers to a certain city, and that "x" refers to an unknown quantity. These are not referents that we would learn ourselves through experience with the perceived qualities of objects, or experience acting on objects. When Bruner talks about symbolic systems, he talks about language, and it is clear that he considers language the essential symbol system. Linguistic representation is the mode which is most interesting to Bruner in terms of its implications for cognition. However, since this discussion is concerned with non-linguistic representation in early cognitive functioning, we will focus on the stage when iconic processing is predominant.

Bruner discusses various studies which demonstrated the salience of perceptual cues for young children, and he argues that these results provide evidence for the predominance of iconic representation at certain ages. One study examined comparative judgments of fullness with children 5, 6, and 7. The children were shown pairs of glasses containing water, and they were asked which glass was fuller. In the various pairs the size of the two glasses was sometimes the same and sometimes different, and the proportion of water in the two glasses was sometimes the same and someimes different. The children almost always chose the glass with the greater volume of water as the fuller, even when the other glass was filled to the brim. This finding demonstrated the salience of visual information for young children. Lawson (1978) showed that children are likely to use a perceptual strategy when a task becomes complicated. She asked children to make equivalence judgments about rows of dots and found that some young children were consistently influenced by length of row and others were consistently influenced by number of dots in the row. When a large number of dots were used in the stimulus materials, making the task more difficult with regard to numerosity, then many more of the children relied on a perceptually-based length strategy.

The importance of perceptual information for children has also been shown by Francoise Frank (cited in Bruner, 1964). She compared a standard conservation task in which water is poured from a standard beaker into a wider beaker, to the same task with the pouring done behind a screen. All the children received both tasks, and in each case they were asked whether the amount of water was the same, greater, or less, after the transformation. In the classic conservation task four- and five-year-olds performed very poorly, and six- and seven-year-olds made the correct response about half the time. When the same children were shown the beakers before and after the transformation, but they could not see the water actually being poured, performance improved dramatically. The four- and five-year-olds gave correct answers fifty percent of the time, and performance was perfect for the six- and seven-year-olds. These results show the significance of perceptual information for young children, and implicate perceptually-based representation in the early years.

In another approach to studying early reliance on iconic process, Bruner reports that children were given matrix problems in which items varied along two dimensions. Children 5, 6, and 7 were able to copy the matrix designs with no difficulty. But when the matrix was to be transposed from the demonstration matrix, that is, reversed or rotated, the five- and six-year-olds had great difficulty. Most of the seven-year-olds were able to do the task. What is important to note is that the the seven-year-olds talked to themselves about the problem. Bruner concluded that the younger children were more fixed to the original image presented, and were not able to attend to the principles of organization in the matrix. The seven-year-olds evidentally used language to transpose the matrix in a way consistent with the principles of organization.

Eleanor Rosch has made important contributions to the study of classification, for example in describing natural levels of categorization. Rosch, Mervis, Gray, Johnson, and Boyes-Braem (1976) presented evidence that there is a basic level of categories which is first learned by children. Basic level categories have maximum cue validity so that information is more easily attained than for other category levels. Objects in basic categories, for example tables and chairs, have more attributes that distinguish between categories than do objects in lower category levels, such as kitchen chairs and dining room chairs. And, basic level objects have more common attributes within their category than do objects in higher category levels, like furniture and vehicles.

This work has some bearing on the question of non-linguistic representation. Rosch (1977) comments that most work done on categorization assumes a feature analytic model in which a small number of discrete criterial attributes provide the basis for classification. In this research approach, the materials used are well suited to verbal handling, and are not well suited to nonverbal processes. Because of this bias in experimental approach, she argues that much past research did not reasonably describe natural classification functioning. Rosch proposes, instead, an analog process in which items are compared to a prototype, and are more or less good matches to that prototype. Categories consist of attribute domains clustered around a prototype which is perceptually salient, or which possesses maximum cue validity, providing maximal information about the category.

Similar to Bruner's explanation of the basis of iconic representation, Rosch believes that prototypes are generally constructed from experience with category instances, and so there is a developmental aspect to category acquisition. While the real world is organized in a way that affects categorization, our perceptual and information-processing systems affect how we interact with the world and construct category prototypes. It is not clear in Rosch's writings exactly how prototypes are represented. Rosch suggested that prototypes may be defined rather broadly

> for example, as the abstract representation of a category, or as those category members to which subjects compare items when judging membership, or as

> the internal structure of the category defined by subjects' judgments of the degree to which members fit their "idea or image" of the category (Rosch and Mervis, 1975, p. 575).

For perceptually-based categories and categories at the basic level of natural classification, Rosch's discussion suggests that the prototypes may be perceptually-defined, and could be imaged. If this is the case, and if visual or iconic representation is the preferred system for young children, it is understandable that basic level categories are the first categories learned by children. Basic level categories are more easily represented by a visual image than are higher levels of natural categories, and the images to represent basic level categories are more distinctive than is the case for lower category levels.

Another psychologist who has discussed non-linguistic representation is Lee Brooks. Like Rosch, Brooks (1976) points out that most previous research has been structured to force subjects to analyse critical attributes, to extract the criterial generalities of classes, and he also concludes that this approach does not capture all of natural classification functioning. While Bruner and Rosch are talking about category representation with generalized images, or prototypes, Brooks proposes a more specific process, matching to instance. He argues that when people are confronted with a novel instance to evaluate, they attempt to evaluate it on the basis of its physical similarity to previous instances. This is an analog process in which the novel item is matched to the known item which it most resembles, that to which it is perceptually most analogous.

Brooks conducted a series of studies in which university students were either instructed to learn to discriminate categories, or were given a paired associate task with the same materials. He assumed that with the first instructions subjects would look for criterial attributes, while in the second case they would not do so. The materials used ranged from artificial grammers embodied in English letters to drawings of cartoon animals against different backgrounds. After the first part of the session, category learning or paired associate learning, all subjects were tested for ability to categorize new items. The subjects in the paired associate condition categorized new items better than the subjects who had received categorization instructions and experience. Brooks concluded that these successful subjects were matching to the learned instances, while the subjects in the other condition were trying to use analytic strategies because of the category discrimination instructions. Matching-to-instance was a more successful strategy with the complex and unfamiliar materials used.

Based on these findings, Brooks concluded that much classification functioning may involve matching to best instance, a process that is perceptually- and not linguistically-based. Moreover, matching-to-instance appears to be an appropriate and successful process with certain kinds of tasks. These results show that perceptually-based representation persists into adulthood.

Representation May Depend on Stimulus Content

The mode of representation, and the relationship between language and classification, may depend upon the nature of the stimulus content. Piagetian research, and the research I have done, show that the age of logical competence varies from one study to another, even when identical procedures are used. Subject sample characteristics may account for some of this variance, but it appears that experimental materials are also responsible for differences obtained. Stimulus materials may vary in terms of type (e.g. geometric forms, pictures of real life objects), complexity (number of potential distinguishing attributes), and hierarchical category level (e.g. collies, dogs, mammals, animals). Just as age of competence varies with stimulus content, so the relationship of language and classification may depend on the content of the classes.

For example in the case of category level, Rosch (1976) has demonstrated that there is a natural basic level of categorization. This basic level is the first one learned by children, and verbal labels are learned at an earlier age for the basic level than for the other levels of categories. Consideration of the examples Rosch gives for the lower, basic, and higher levels of categories suggests that compared to other category levels basic level classes are more likely to have a one-word name, or a name which is familiar to children. Examples for lower, basic, and higher category levels are: bald eagles, birds, things that fly; station wagons, cars, things people ride in.

Another aspect of stimulus content which may affect the relationship of language and classification is modality. Blank (1974) points out that most cognitive developmental research uses material which is predominantly visual and this may affect our understanding of the role of language in real-life cognitive development. Blank argues that children's earlier and preferred method of communication is gestural and visual. Thus when children are presented with visual material language may be an impediment, because language must then compete with a better established system of representation and communication.

In support of this position, Blank reported data from studies which examined language and modality of the task. When three-year-olds were holding an object in each hand, they were almost always able to follow a verbal instruction to drop the one in their right hand. When the same instruction was given for visually displayed materials ("Point to the one on the right") the children usually ignored the instruction (Blank and Altman, 1968). Blank also found that language helped children on tasks involving temporal concepts. Nursery school children were proficient at discriminating between one and two circles whether they used language or not. But when required to distinguish between one or two successive light flashes, these children could succeed only if they used appropriate language. If they were taught verbal labels, they were able to do the task (Blank and Bridger, 1964). Oleron (1957) obtained similar results when he studied transposition learning in deaf children. He compared spatial and temporal material, and found that the children learned transposition with the spatially-defined materials, but failed to do so when the material was based on temporal concepts such as fast and slow. Since deaf children are deficient in langauge, Blank considers this evidence that language is not essential for visual materials, but that language is useful in dealing with non-visual materials.

Blank's findings concerning modality and use of language may be relevant to classification tasks. Language may be important for the classification of sounds, textures, odors, or items which vary on a temporal dimension. But much of what we classify in real life, and virtually all of what we classify in experiments, is defined in the visual modality. Thus, Blank would predict that language is not particularly useful for young children in most classification.

Chapter 3

Class Inclusion

Piagetians consider classification and seriation the two basic intellectual functions in the concrete operational stage of cognitive development. Seriation is ordering items along some quantitative dimension, such as light to dark or long to short. Classification is the recognition, construction, and ordering of classes. Of the two areas of cognitive operation, classification raises more complex issues (Inhelder and Piaget, 1964/1969, p. 1), and the work reported in this thesis concerns one classification operation, class inclusion.

A class is a set of items which have a common attribute, and classification is based on evaluation of common and distinctive attributes. In hierarchical classification there is a structure such that some classes are combined into other classes, for example, collies and boxers are dogs, dogs and cats are animals. Obviously most of the classification that we deal with is hierarchical in nature. When children are competent in hierarchical classification they understand the relationships between different levels of the class structure, so that they are able to relate the parts to the other parts, the parts to the whole, and the whole to the parts.

Piagetians assert that mature understanding of the relationships between different levels of the hierarchical classification structure is best assessed by the class inclusion question in which the child is asked to compare the size of a sub-class and its superordinate class.

> The conservation of the whole and the quantitative comparison of whole and part are the two essential characteristics of genuine class-inclusion....it is the basis of any classification which really does order classes, instead of only differentiating between them. (Inhelder and Piaget, 1964/1969, pp. 117-118)

Given two subclasses, A and A', which are part of the superordinate class, B, the class inclusion question requires quantitative comparison of the larger subclass, A, and the superordinate, B: Are there more A or more B? For example, given three dogs and two cats, the inclusion question is: Are there more dogs or more animals?

Development of Classification Function

Piagetians have extensively investigated many aspects of classification operation. Children's production ability has been tested by asking them to sort items, usually by saying, "Put the things that are alike together," or some variation of this request. The first kind of sorting children produce, between the ages of 2 and 5, consists of "figural" or "graphic collections", in which children put items together to form spatial patterns or figures. Around age 5-1/2 children produce "alignments", by matching items one to another on the basis of some perceptual similarity. This matching is done in a chain sequence, and the basis for matching changes as the children's attention shifts from one perceptual attribute to another. This sorting is not consistent or stable. Between 5 and 7, children's classification ability matures, and sorting becomes more systematic. Children are now able to plan ahead and employ an overall scheme for sorting, without switching the basis for sorting in the arbitrary manner which characterizes "alignments". The sorting is exhaustive, with all items being put in the appropriate classes. In time children can construct hierarchical classes, progressing from subclasses to superordinate grouping, or the reverse, sorting to superordinate classes and then subdividing. During this period children attain greater flexibility in sorting, and can reclassify items in different ways when asked to do so. In one study with items which varied in shape, color, and size, 76% of the six-year-olds were able to successively sort using two or three criteria. Multiple classification, in which children demonstrate ability to cross-classify on two attributes, is mastered around age 7 or 8.

According to the Piagetians, the final component of classification function is class inclusion, the ability to quantitatively compare subclass and superordinate class. Inhelder and Piaget (1964/1969) found that inclusion competence is exhibited between 8 and 12 years of age.

Kofsky (1968) conducted a large-scale study to examine the sequence of classification abilities described by the Piagetians. While her tasks differed to some extent from Piagetian procedure, overall the results obtained fit the normative information reported above. Two tasks, resemblance sorting and consistent sorting, had requirements like "alignment" sorting; the children had to match a stimulus item, and group at least three like items. Over 80% of the four-year-olds succeeded at these tasks. Seventy-five percent of the five-year-olds were able to do exhaustive sorting. Sixty percent of the six-year-olds were able to answer questions about multiple classification, and two-thirds of the seven-year-olds reclassified items. The order of attainment of multiple classificaton and reclassification abilities is the reverse of that reported by Piaget; however, this may be an effect of the procedures or materials Kofsky used. There is no task in the study analogous to hierarchical sorting.

On class inclusion, Kofsky reported that 60% of the nine-year-olds passed the criterion on the inclusion task. The criterion was two correct out of three answers, which is easier than the usual Piagetian requirement of consistent correct answers accompanied by explanations of those answers. In addition, only two of Kofsky's questions were Piagetian inclusion questions. The third question did <u>not</u> ask for an inclusion comparison. However, the conclusion remains the same: class inclusion is the last classification ability demonstrated.

The sequence of classification skill acquisition is one basis for the Piagetian assertion that classification mastery is best measured by class inclusion. Because children can do hierarchical sorting before they can correctly answer the inclusion question, it would seem that ability to recognize and sort hierarchical classes does not indicate full understanding of the relationships between different levels of classes. Piagetians point out that to successfully answer the inclusion question children must <u>simultaneously</u> consider the sub-class and the supra-class. They must have access to the extension of B and the extension of A at the same time in order to compare the sizes of the two. But successful hierarchical sorting requires only <u>sequential</u> consideration of the various levels of classes. For example, given a set of red triangles, blue triangles and yellow squares, children may sort first into red, blue, and yellow groups. Then, they can ignore color and focus on shape, so that triangles are now put together and distinguished from squares. The point is that the consideration of the two different criteria, color and shape, need not occur at the same time for this sorting.

Another argument made for class inclusion as the definitive measure of classification ability is based on the errors that young children make. Before inclusion competence is available, children tend to systematically give the wrong answer, so that percentage of correct responses for certain age groups is below 50%. It appears that this is so because the children are comparing the subclasses, A and A', rather than comparing subclass and superordinate class, A and B. (This explains why the larger subclass is always the one used in the inclusion question. If children are using subclass comparison, there may be false correct response to the comparison of A' and B. The children may say B is larger, meaning, in effect, that A' is smaller than A.) Piagetians conclude that when inclusion competence is not available because children cannot simultaneously compare superordinate class and included subclass, the children resort to disjunctive comparison of the two subclasses. In other words, the erroneous strategy of comparing subclasses to each other is considered by Piagetians to result from logical deficit. However, class inclusion studies done by researchers outside the Piagetian school indicate that logical deficit or competence is not the only factor which affects class inclusion performance. The results obtained on an inclusion task are affected by the materials used, perceptual factors, and the language employed.

Perceptual Effects

Smedslund (1964) used geometric figures for items in the inclusion tasks. For example, children 4 through 11 were shown red and white squares and circles, and were asked certain preliminary questions. Then the white items were removed, leaving ten red circles and three red squares. The remaining items were covered up, and the children were asked the inclusion comparison, whether there were more circles or more red things. Finally the items were uncovered, and the inclusion question was repeated. One finding of interest is that with these simple materials, varying only in color and shape, inclusion competence was demonstrated at a relatively young age: about 90% of the seven-year-olds made the correct inclusion response. The other aspect of this procedure which is important concerns asking the question first with the items covered, and second with the items in view. This made little difference with the older children, but with the younger age group (4:3 to 6:2) almost 15% of the children answered correctly when the items were concealed and then reversed their answer when the items were exposed. Since some of these younger children answered the inclusion question better when the items were covered, it appears that perceptual display, that is, having items visually available, does not facilitate correct response to the inclusion question. Unfortunately, there is an alternate explanation because of the methodology used. The younger children may have tended to reverse their answers because the uncovered-item-question was always asked second, and they may have assumed the question was repeated because their first response was not right.

Wohlwill (1968) also compared inclusion response when the items were present and when the items were not present. Half of the children received the items-present questions first, and half received the items-absent questions first. The children gave more correct responses when they could not see the display. Wohlwill concluded that the perceptual information predisposed the children to focus on the subclasses, not on the superordinate class as an entity, and so they were more likely to compare the subclasses to each other. However, neither Brainerd and Kaszor (1974) nor Winer (1974) supported Wohlwill's interpretation. These authors concluded that the effect found by Wohlwill may have resulted from stating the number of items in classes for the items-absent questions, but not for the items-present questions.

Wohlwill also investigated the question of perceptual salience by adding items which were not in either subclass, and therefore not in the superordinate class. The idea was that the contrasting items would increase the salience of the superordinate class so that children would be more likely to make the correct response, and performance was in fact better when the additional class was used. Isen, Riley, Tucker and Trabasso (1975) pointed out that Wohlwill did not give transfer tests to single class problems, that is, problems which did not contain the contrasting items outside of the superordinate class. It is possible that better performance resulted not from increased salience of the superordinate class but from a perceptual strategy of comparing the larger subclass to all of the other items, the items in the smaller subclass and the contrasting items outside of the superordinate class. If so, the children were not making an inclusive comparison. Consequently Isen et al. ran a procedure similar to Wohlwill's and tested for transfer to single class problems. The improved performance obtained with the contrasting items transferred on the post-tests, and thus it appears that perceptual cues which increase the salience of the superordinate class facilitate children's inclusion performance.

Isen et al. used two ratios of larger to smaller subclass, 6:2 and 4:4, and found that children gave more correct answers with the smaller ratio, 4:4. This result replicated the findings of Ahr and Youniss (1970) that as the larger subclass (A) is relatively bigger than the smaller subclass (A'), children are more likely to make the incorrect response, saying that A is larger than B, the superordinate class. One could criticize the Isen et al. finding on the grounds that with subset sizes of 4 and 4 the children were not being asked an inclusion question, since the inclusion question is defined as the comparison of _larger_ subclass to superordinate class. However the Isen et al. results are consistent with those of Ahr and Youniss who used the same two ratios and three other ratios. In the Ahr and Youniss study every child received inclusion questions for subclass compositions 8:0, 7:1, 6:2, 5:3, and 4:4, and with these five ratios percentage of correct response increased as the subclass ratio decreased. Ahr and Youniss concluded that subset ratios affect the likelihood that young children will focus on the subclasses, and Isen et al. agreed with this position. Whereas Piagetians claim that subclass comparison results from logical deficit, these authors concluded that use of that strategy is also affected by the perceptual context, in this case the relative size of the subclasses.

Brainerd and Kaszor (1974) disagreed with a perceptual set interpretation of the subclass ratio findings. They found that children were likely to give correct answers for equal ratio sets (4:4) to the usual inclusion question about difference, "...more As or more Bs?" But for unequal ratio sets (5:3, 6:2, 7:1), children were likely to give correct answers to an equivalence question, "Are there the same number of As and Bs?" In both of these cases, Brainerd and Kaszor concluded that the correct answers contained some "false positive" response. With equivalence questions for unequal ratio sets, children correctly reply that A and B are not the same whether they are comparing A and B, or comparing the subclasses, A and A'. But with the difference question, if children are comparing subclasses of equal size they are likely to guess, which leads to correct answers half the time. Thus, the subset ratio effect is not a perceptual effect, but is an artifact resulting from the combination of question form (equivalence, difference) and ratio type (equal, unequal). Brainerd and Kaszor's results indicated that propensity to compare the subclasses was not affected by subclass ratio per se. This conclusion is consistent with the Piagetian view that comparison of subclasses is a strategy based on logical deficit.

Wilkinson (1976) manipulated the salience of the superordinate class by referring to distinctive perceptual features of that class. Forty-eight pre-school children were given a standard inclusion question, and an inclusion question which included distinctive features of both the subclass and the superordinate class. For example, the children were shown line drawings of adult figures, all of whom had a chair; some of the adults were women with a picnic basket, and the rest were men without a picnic basket. The standard inclusion question was "Are there more mothers or more grown-ups?", and the distinctive features question was "Are there more grown-ups who have a picnic basket or more grown-ups who have a chair?" With the standard question, children gave the correct answer 23% of the time, but when specific identifying features were used for the classes the children were correct 60% of the time.

Wilkinson offered a counting strategy explanation of his results. He suggested that the first counting strategy available to children includes a prohibition against counting items twice; thus children count A and then B - A = A', which results in comparison of the subclasses. Later counting strategy allows for A and B to be counted separately, so that the A items are counted twice, first in A and then as part of B, which results in the correct comparison. His use of perceptually distinctive features for the classes allows young children a way around the double-counting prohibition of the first strategy. In the example above the children could count picnic baskets and chairs, rather than double-counting the mothers. Or, in another case, Wilkinson compared a "concept" question, "Are there more boys or more children?" to a "percept" question, "Are there more houses with doors or more houses with windows?," where all of the houses had windows, but only some of them had doors. Again, the children could count the windows and then the doors, whereas they could not count boys separately from children. The five-year-olds gave a correct inclusion answer more often for the house question than for the children question.

Wilkinson's work indicates one way in which the perceptual characteristics of stimulus sets may affect inclusion performance; given an immature counting strategy, children are able to quantify the superordinate when that class has distinctive features. However, there are at least two other possible interpretations of these findings. It may be that the use of distinctive features in identifying the superordinate class makes clear to the child the comparison which is requested. In the standard question, for example mothers and grown-ups, the child may assume the experimenter is asking for the comparison of mothers and other grown-ups (this comprehension problem will be discussed in detail later). In contrast, for the "percept" questions, it would be clear that the experimenter is asking for comparison of subclass and superordinate, "with picnic baskets" and "in a chair". This hypothesis does not conflict with the double-counting explanation. It is possible that both factors, understanding of the request and counting strategy, are contributing to the better results with the "percept" sets.

The other interpretation derives from the operationalization of the "percept" sets, and it raises a serious possible objection to Wilkinson's conclusions. Because there are separate and distinctive features defining the subclass and the superordinate class, one could argue that children are in fact treating these as exclusive classes and not as subclass and superordinate class. For example with the houses, it may be that the children are not counting houses with doors and houses with windows, but are simply counting doors and windows. If this is the case, the children are not making an inclusive comparison, but are comparing disjunctive classes. Thus, providing distinctive perceptual features for the superordinate class would not be facilitating inclusion performance but rather removing the requirement to make the simultaneous comparison of subclass and superordinate class that is basic to Piagetian inclusion.

Tatarsky (1974) posited that the dimensional salience of the attributes which define the subclass and superordinate class may affect inclusion performance, and he proposed that typically the attribute for the subclass is more salient than that for the superordinate class. To test this hypothesis, he gave children aged 5, 6, 7 and 8 three different kinds of materials, and asked inclusion questions about these materials. All of the children received "two-dimension, unequal salience" questions about wooden cylinders which were painted all blue or all red: "Are there more blue blocks or more wooden blocks?" These questions represented the standard Piagetian condition. Half of the children were asked "two-dimension, equal salience" questions about cylinders which were half painted, either blue or red, so that the other half of the cylinders was unpainted wood. The question was the same, "...more blue blocks or more wooden blocks?", but in this case the superordinate attribute, "wooden", could be seen as readily as the subclass attribute, "blue". The other children received "one-dimension" questions about cylinders which were half yellow, with the other half painted either blue or red: "...more blue blocks or more yellow blocks?". In this case subclass and superordinate attributes were not only equally visible, or salient, but these attributes were from the same dimension, color. The children gave more correct inclusion responses to the one-dimension questions than to the two-dimension, equal salience questions, and performance was better for two-dimension, equal salience than for the standard Piagetian case with two dimensions which are unequal in salience.

The materials that Tatarsky used have a perceptual basis similar to those used in the later study by Wilkinson, and the observations made about Wilkinson's "percept" conditions apply to Tatarsky's two "equal salience" conditions. The improved performance in the two-dimension, equal salience and the one-dimension questions may result from children understanding the requested comparison more clearly; or it may result from children treating the experimental subclass and superordinate class as two separate classes. Tatarsky's finding that the children did better in the one dimension case than in the two dimension-equal salience case weakens Wilkinson's argument about counting strategy. If use of an early counting strategy which prohibits double counting were the only basis for improved performance when the superordinate attribute is more salient, or more countable, then there should be no difference between Tatarsky's two "equal salience" conditions. The superordinate class in these two cases would have been just as countable for the young children regardless of whether they were counting wooden cylinder halves or yellow cylinder halves.

McGarrigle, Grieve, and Hughes (in press) manipulated the salience of the superordinate class in a manner very similar to that of Tatarsky and Wilkinson, and obtained similar results. Using black cows and a white cow, all lying down, they compared the standard Piagetian question, "Are there more black cows or more cows?", to the expanded form, "Are there more black cows or more sleeping cows?" For children 3 to 5, superordinate responses occurred 31% of the time with the standard question and 52% of the time with the expanded question. For children 5 to 7, correct response went from 25% on the standard question to 43% on the expanded form.

McGarrigle et al. also investigated the effect of increasing the salience of the subclass, predicting that this would increase use of subclass comparison. With one set of materials a teddy bear was four red chips ("steps") from a chair. Two white chips beyond the chair was a table, so that the table was six chips from the teddy bear. When the subclass was not salient, the question was analogous to the Piagetian inclusion question: "Are there more steps to the chair or more steps to the table?" When the subclass was made salient, the children were asked, "Are there more red steps to go to the chair or more steps to go to the table?" When the subclass was not salient, performance was better than that typically obtained with the standard Piagetian question. Children 3 through 5 gave correct answers 64% of the time. This suggests that either this question was more interesting to the children than "Are there more red chips or more chips?", or, the meaning of the question was somehow clearer to the children than the Piagetian question. However, when the subclass was more salient (red steps or steps to table), the children's performance dropped to 38% correct responses. Thus, McGarrigle et al. showed that inclusion performance improves when the superordinate class is made more salient, and is worse when the subclass is made more salient.

The investigation of superordinate class salience in the inclusion question has been taken a step further by Markman (1973) and Markman and Siebert (1976). They have shown that children in kindergarten and first grade perform better on inclusion questions when the superordinate class is labeled with a collective noun like family or pile, than when a class noun like dogs or blocks is used. In discussing this finding, Markman and Seibert point out certain key distinctions between collections and classes. Collections are defined by the relationship of the members (parent dogs and baby dogs make a family of dogs), while classes are defined by a common attribute possessed by each member ("dogginess"). In a collection, the parts are not representative of the whole as is true for a class; baby dogs are not a family of dogs, but they are dogs. Finally, collections have an internal organization which is lacking in classes. Parts fitted together make a machine; grapes attached to a stem form a bunch; blocks arranged one on top of the other are a pile. Classes, on the other hand, are aggregates of items which have a common attribute, or set of attributes, but the classes do not have any internal structure. Because of these differences, Markman and Seibert conclude that collections are more coherent psychological units than classes. This greater "wholeness" on the part of collections allows children to preserve the superordinate class better when a collective noun is used in the inclusion question, and this accounts for the better performance which is observed.

Turning to another aspect of perceptual effects and class inclusion ability, Markman (1978) provided evidence that even when children answer inclusion questions correctly and can justify their answers, the basis for performance may be empirical rather than logical. Subjects were children aged 7, 8, 9, 10, and 11 who succeeded with Piagetian inclusion questions. These children were asked various questions designed to measure knowledge of the part-whole relationship when the classes were not perceptually available. For example, the stimulus items were put behind a screen and the experimenter said, "...I am going to take some of the furniture away...can you tell for sure whether there is more furniture left or more couches left?" In another type of question the children were asked, "Could you make it so that there will be more spoons than silverware on the table?" The children younger than age 11 had difficulty with these experimental questions. Since the young children could not demonstrate inclusion logic when empirical information was missing, the authors concluded that the children's success with the Piagetian inclusion question was not based on understanding of logical necessity.

Markman's findings are not incompatible with the Piagetian position. Piagetians are aware of discontinuities in the application of children's understanding, and in fact say that when children are in the stage of concrete operations, around ages 7 through 11, they can apply logical operations to concrete objects but cannot apply the same operations with abstractions. The limitations in dealing with abstractions imply that logical development is not fully mature, but nonetheless the children are dealing with concrete objects logically. Also, it should be noted that Markman's procedure added transformation of stimulus materials to the standard inclusion problem, and probably the transformation made the problem more difficult for the children. It would be useful to compare Markman's results with a condition in which the same question is asked with the stimulus items displayed; for example, with silverware on the table the children could be asked, "Can you make it so that there will be more spoons than silverware?" The children might do just as poorly in this case as when the materials are concealed. If so, the difficulty results from the added problem of the transformation, not from the lack of empirical information.

The inquiry into perceptual effects on class inclusion has been added to by Isen, Riley, Tucker and Trabasso (1975) and Trabasso, Isen, Dolecki, McLanahan, Riley and Tucker (in press) in a series of experiments. As already mentioned, in one procedure these researchers replicated Wohlwill's finding that use of contrasting items (D) outside of the superordinate class facilitates inclusion performance, presumably because the superordinate class (B) is more salient to the children because of the contrast provided by D. In the Isen et al. study the D class was composed of two subclasses, C and C', with equal numbers of items in these two subclasses. An example of a stimulus set would be four apples (C) and four oranges (C') making eight fruit (D), and six dogs (A) and two cats (A') making eight animals (B). The children gave more right answers to the inclusion question, "Are there more dogs or more animals?", when the fruit was present than they did when only the animals were displayed.

In a second procedure in the Trabasso studies, some of the children were asked all pair-wise contrasts for the two superordinate classes, for example "more dogs or more apples," "more dogs or more fruit," "more animals or more fruit," and so on. These questions included the class inclusion question for both B and D, that is "more dogs or more animals," and "more apples or more fruit." This use of multiple comparisons facilitated inclusion performance in children 5, 7 and 9 in comparison to single class (standard) inclusion questions and double class questions which did not include multiple contrasts.

Even more striking than the effects with a contrasting class and multiple contrasts, is the evidence obtained from the comparison of a subclass to an unrelated superordinate level class, for example, "more dogs or more fruit" (A r D). In both of the Isen et al. studies there was no difference in correct response when children were asked, "Are there more dogs or more animals?", or when they were asked, "Are there more dogs or more fruit?" This is an astounding finding. Piagetian theory would not predict difficulty with the A r D comparison, because the two classes are distinct and there should be no problem with simultaneous consideration. Wilkinson would predict that the A r D comparison should be easier than the inclusion question because items do not have to be counted twice in the between class comparison. An information processing analysis presented by Klahr and Wallace (1976) also would predict that the between class comparison is easier than the inclusion comparison for young children, because children should correctly encode quantity information about the superordinate class D which is identified as "fruit-not-dogs". The task analysis provided by Trabasso et al. does account for the finding of equivalent difficulty for the two types of questions. This model proposes that the children, given the display with the two superordinate classes animals and fruit (B and D), encode only the subordinate classes, dogs (A), cats (A'), apples (C) and oranges (C'). Thus when asked to compare dogs (A) and fruit (D), the children do not have an accurate estimation of fruit, (D), but must use either apples (C) or oranges (C'), which results in incorrect response.

In addition to the study of perceptual effects, the findings of Isen et al. and Trabasso et al. have implications with regard to language. Isen et al. (1975) speculated that the child

> "has to figure out what the speaker _means_ by the question...language, then, when viewed as communication, is shown to be highly context dependent for children and the interaction of language and perception in class-inclusion contexts leads to a misinterpretation of the question and a subsequent failure to employ the correct operations (pp. 8, 9).

<u>Language Effects</u>

In the studies just discussed, perceptual factors were manipulated and this affected inclusion performance. It was pointed out that in some of these studies performance may have improved not only because of perceptual effects but also because the children understood the question better. An example is a study by Sheppard (1973). With a standard inclusion set, B = A + A', six-year-olds were given training in which the larger subclass, A, had a matching class, A_1. The training procedure consisted of showing the children that A was equal in number to A_1, B equaled $A_1 + A'$ which was greater than A, and therefore B was greater than A. These relationships were discussed with the children and they were "guided to see the relationships". The training was effective, and improved performance on the inclusion question was found a week, two to three weeks, and three to four months later. The manipulation of classes, and especially the discussion of the manipulation, may have made clear to the children what the class inclusion question means.

This discussion of question comprehensibility assumes that the Piagetian form of the inclusion question is difficult to understand. Various writers (e.g., Markman, 1973; Kalil, Youssef & Lerner, 1974; Trabasso et al., in press) have suggested that the conjunction "or" may be misleading because this word usually joins two disjunctive, or separate, classes. There is evidence that the Piagetian question implies disjunctive comparison even for adults. Winer (1974) asked college students, "In the whole world are there more dogs or more animals?", and then asked them whether or not they assumed animals in this question included dogs. Thirty-one percent assumed that the comparison was between dogs and other animals, that is between the specified subclass and the superordinate class minus that subclass. Winer's conclusion seems quite reasonable: if that many adults misunderstand the question, probably a great many children misunderstand the question.

In an attempt to clarify the inclusion request for children, Winer altered the standard Piagetian inclusion question to a more elaborate form: "If I had four apples and three pears, would I have more apples or more things to eat?" He found that children of 7, 8 and 9 gave more right answers to this question than to the usual question, "Are there more apples or more things to eat?", and he concluded that with the elaborated verbal form children were better able to understand the logical comparison which was requested.

Ahr and Youniss (1970) also used an expanded verbal form of the inclusion question in a training study, along with a simple verbal feedback procedure. One group of children received verbal feedback only, and the second group was given verbal feedback plus an expanded form of the inclusion question: "Are there more B or more A or more A'?" Both groups of children showed improved performance during training, but the children who received the expanded verbal form did not maintain higher performance in post-testing. In the verbal feedback only condition, however, higher performance was maintained: 60% of the eight-year-olds and 20% of the six-year-olds learned successful inclusion performance. The authors concluded about the success of the feedback only training that "...they had understood inclusion but failed to perform appropriately. This argument, compatible with Wohlwill's (1968), recommends a distinction between comprehension of an operation and the expression of it at the level of performance... (p.142)" I suggest that the children became able to perform, and to demonstrate their competence, because they learned from verbal feedback what the experimenter was asking for in the standard Piagetian question. In the expanded form of the question which specified both subclasses and the superordinate class, the intended comparison was immediately clear to the children. However, in transfer testing, the original non-expanded Piagetian form was used and the advantage of understanding the question was lost.

Kohnstamm (1968) used not only feedback but also detailed explanation and discussion to train five-year-olds. This verbal training was effective: in three experimental groups, between 30% and 80% of the children showed improvement in class inclusion performance.

Digressing from the discussion of language effects, there was an interesting finding in Kohmstamm's study about type of stimulus material. He used three presentation conditions: in his first group classes were presented verbally, in the second group, pictures of objects were used to illustrate classes, and in the third group, pictures of blocks were used. In the first two groups, 30% and 40% of the children achieved class inclusion. But in the third group, although the procedure was the same, 80% of the children were able to answer class inclusion questions, and performance was even better at the three-week post-test. The difference in this last case was in the stimulus materials used. The author's only comment on this was, "...it was apparent that the block material had a much greater learning effect than the picture material...". It appears that the simplest stimulus material was most effective in Kohnstamm's training, and this is especially interesting because the transfer tests for this group included questions that were purely verbal, and questions about pictured material.

Like Ahr and Youniss, and Kohnstamm, Brainerd (1974) examined verbal training. He used simple verbal feedback in attempts to train children 4 and 5 in transitivity, conservation, and class inclusion.

He reported that it was more difficult to train the children for inclusion than for transitivity and conservation, but the inclusion training procedure was somewhat effective. There was no correct response in pretesting; in the post-testing, one week later, the children gave correct answers 41% of the time.

The Kohnstamm, Ahr and Youniss, and Brainerd results are consistent: there was some improvement in inclusion performance following verbal feedback training, and that improvement was maintained in later testing. All of these authors concluded that the training was effective, that the children were learning to make the inclusion comparison. However, it could be that feedback was effective because it made clear to the children what the inclusion question meant. Of course, training effects and question clarification may have jointly contributed to improved performance.

Siegel, McCabe, Brand and Matthews (1977) studied language effects on inclusion performance by varying the linguistic form of the inclusion question. They asked children, "Do you want to eat the smarties or the candies?", and compared this to the standard Piagetian form, "Are there more smarties or more candies?" Three- and four-year-olds were more likely to know what they wanted to eat than to correctly state which class had more. In other words they gave correct inclusion response more often to Siegel's question than to the Piagetian question. These performance differences may be based on the language used, "do you want..." versus "are there more..." In addition, there probably are motivational effects because young children have more personal interest in what they eat than in quantitative comparisons.

Assessment of Inclusion Ability

The research surveyed here clearly indicates that when we study class inclusion we are not assessing only inclusion competence. There are many factors which affect children's response to inclusion questions. The context of the questions is important, and this includes the kind of class content used, the salience of the attributes which define subclasses and superordinate class, and the presence or absence of contrasting classes. Quantification skills available to the children, and the use of multiple comparisons may affect response. Linguistic factors are significant, especially with regard to whether or not the children understand what it is that we are asking them to do. If North American adults were asked the inclusion questions in Greek, most of us would fail the question.

Chapter 4

Purpose of Thesis

Overview

In the assessment of class inclusion reported in this thesis, language was varied in two ways. First, different forms of language were used for the inclusion question; each linguistic form posed the same logical demand, quantitative comparison of subclass and superordinate class. Second, language was reduced in the standard Piagetian inclusion question. One purpose of treating language as an independent variable was to examine the sources of error in class inclusion performance in order to determine how much error is language-related, and to identify the circumstances under which language-related error occurs. Another purpose of using different language conditions was to obtain a more accurate description of developing inclusion ability through a research approach of converging estimates.

One product of this work serves both of these goals. For the last experiment a mathematical model was developed which partitions class inclusion response into estimates of the underlying components of that response. Using this analysis it is possible to estimate the extent to which children are using inclusion logic, or comparing the subclasses, or guessing. The only data used by the model are children's answers to comparative questions; there is no analysis of children's explanations of their answers.

The mathematical model is a useful tool in the study of class inclusion because it allows examination of the strategies, and hence the logical bases, used by children in different assessment conditions, and at different ages. Thus, in a language-defined assessment condition it is possible to determine how much error is based on comparison of subclasses which always leads to the wrong answer, and how much error reflects some kind of guessing strategy, which leads to the wrong answer about fifty percent of the time. Similarly, the correct answers obtained in any assessment condition can be partitioned into those based on inclusion logic and those which represent half of the guessing. The mathematical analysis, then, allows detailed comparison of the effects of different language conditions in assessment, and affords a more precise understanding of developing inclusion competence.

Discussion of Research Methods and Goals

The rationale for comparing different language forms in class inclusion assessment is to provide different estimates of inclusion ability, and to try to understand how the different linguistic forms may account for certain kinds of error. For example, it may be that the Piagetian question is difficult and misleading for children, because the form "Are there more A or more B?" implies comparison of disjunctive rather than inclusive classes. This can be studied by comparing the Piagetian question to an inclusion question form which might be more clear to children. Also, since the Piagetian question asks directly for quantitative comparison of the larger subclass and the superordinate class it may not be a particularly interesting or motivating question for children. One way to test this possibility is to compare the Piagetian question to questions which require the same quantitative comparison, but in a form which is expected to be more interesting for children.

Turning to the reduction of language demands, previous methods used in assessment of conservation, seriation, and transitivity have included techniques based on motivated choice, instrumental choice, and surprise. Many of these studies are variations of a learning paradigm, in which children learn to choose the stimulus item or set which is "bigger", "longer", or "more". The child's choice may be made nonverbally, for example, by pointing. Similarly, with the surprise paradigm the response is nonverbal, and is in fact involuntary: the experimenter judges surprise from the child's facial expression and demeanour. Violation of logical expectancy is inferred from the surprise response. There may be correlative verbal behaviour, but this is not necessary for the surprise measure.

The one study (Braine, 1962) which attempted nonverbal assessment of classification skill used a learning paradigm in which children were to learn to choose either a stimulus set of items that were all in the same subclass, or a stimulus set of items from different subclasses, but all in the same superordinate class. The children were rewarded by a candy hidden under the correct set. There were no verbal instructions about how to choose, and response requirements were nonverbal.

In these previous efforts to assess logical operations with reduced language demand, the stimulus materials were comparable to those used in traditional Piagetian studies, and of course the intention was to require the same logical operation. What is different in these methods is the way of putting the logical demand to the children, and the manner of response (Miller, 1976). In attempting to reduce language in the presentation of the task and in the responses required, I suggest it may be particularly important to consider critical aspects of the logical operation being assessed. For example, in the study of classification, including class inclusion, one critical aspect is the identification of classes. This is basic to classification functioning.

In the standard inclusion question children are asked, "Are there more As or more Bs?", where A and B are the names of the larger subclass and the superordinate class. Thus, the classes in the question are identified verbally. To reduce language, I omitted the names of classes from the standard Piagetian inclusion question, and used visual symbols to identify the classes. With this exception the linguistic form of the inclusion question remained the same, so that this procedure is much closer to the standard Piagetian procedure than is the case in many previous studies which reduced language in the assessment of Piagetian concepts. The similarity of this linguistic form to the Piagetian question is an advantage in asserting that I am testing for the same logical ability as the Piagetians.

Analysis of Response Bases

The mathematical model developed for analysis of components in inclusion response makes it possible to re-evaluate Piagetian interpretation of what children do when they have not yet mastered class inclusion operation. Young children typically give the wrong answer to the inclusion question, rather than giving the right answer half the time as we would expect if they were guessing. The Piagetian explanation is that when children cannot compare the larger subclass to the superordinate class they instead compare the two subclasses, which leads to the systematic error observed. The Piagetian interpretation is reasonable, and it fits the data. However, using only the percentages of right and wrong answers on the inclusion question, it is impossible to know the proportion of responses based on inclusion logic, the proportion based on subclass comparison, and the proportion based on guessing. Piagetians use the clinical interview method to determine the logical bases for children's inclusion judgment. The children are asked to explain _why_ they answered as they did. However, when a logical competence is not yet fully developed, but is in a state of acquisition, it is unlikely that children will be able to explain all of the logical processes which are affecting their functioning.

Brainerd (1973) has written a detailed and influential article about methods of inquiry into the logical bases of cognitive functioning. He argued that there are theoretically-based objections to the use of explanation as a measure of cognitive structure or response strategies. In Piagetian theory, logical operations are internalized actions, such as those involving inversion and reciprocity, and these operations are not based on language. In fact, language is essentially a dependent variable with regard to cognition. Given these assertions of the Piagetian position, use of language to assess logical competence will lead to the Type II error of underestimating logical ability. This is so because all the children who are able to give satisfactory explanations will possess the logic in question, but not all the children who possess the functional logic will be able to explain that logic. Brainerd points out that there is no theoretical reason to assume that use of responses that are simply judgments to evaluate logical competence will systematically lead to either Type II error, or the Type I error of overestimating logical ability.

In this work the mathematical analysis of underlying components of inclusion response uses only children's judgments, the simple verbal responses made to two comparative questions. The data analysed are percentage of correct and error response on the inclusion question and on a question which asks comparison of the smaller subclass and the superordinate class. Because this latter question was not asked in Experiments 1 and 2, the model could not be used with those data. However, the model is very useful in interpreting the effects of assessment conditions in Experiment 3. It is possible to specify why the two language conditions lead to different results because we can determine the strategies which are used in each condition.

To sum up the content of this thesis, the research reported here varied language in the assessment of class inclusion ability with children of different ages. Language was varied by comparing different forms of the inclusion question, and by reducing language in the standard Piagetian question. In the last experiment, mathematical analysis of the components of inclusion performance provided a better estimate than was previously available of inclusion ability and of the other bases for response to the inclusion question.

Chapter 5

Experiment 1: Different Language Forms

This experiment examined the effects of different language forms in assessment of class inclusion. Some linguistic forms may be harder and some linguistic forms may be easier for children to understand. For example, it was suggested in the preceding chapter that the Piagetian inclusion question may mislead children by suggesting disjunctive comparison, the comparison of subclasses. This possibility has been discussed by Shipley (1971), Markman (1973), Winer (1974), and Trabasso, Isen, Dolecki, McLanahan, Riley and Tucker, (in press). To illustrate, suppose the stimulus items are candies, smarties and jelly beans. When asked "Are there more smarties or more candies?", the listener may assume that the intention is to ask whether there are more smarties or more of the _other_ candies. This interpretation, based on linguistic confusion, will lead to the comparison of the two subclasses rather than the requested comparison of the smarties and _all_ of the candies. Winer (1974) has shown that this misunderstanding is common even with adults.

It may also be that the Piagetian question is not particularly interesting or motivating to children. Siegel, McCabe, Brand and Matthews (1977) asked a motivated choice question to test inclusion, and Siegel's question seems more likely to capture children's interest: "Do you want to eat the smarties or the candies?" They reported that three- and four-year-olds gave more correct inclusion response to this question than to the Piagetian question, "Are there more smarties or more candies?" In this study Siegel's inclusion question was one alternate language form which was compared to the Piagetian form. The stimulus materials included candies, marbles, bingo chips, drinking straws, and ribbon. Since most of these materials were not edible, Siegel's question form was changed from "Do you want to eat..." to "Do you want to have...," for example, "Do you want to have the smarties or the candies?" All of the children received four of Piaget's and four of Siegel's inclusion questions to see if they were more likely to answer correctly with Siegel's than with Piaget's question.

It may make a difference to the children whether they receive their more or their less preferred subset in the inclusion question. For example, suppose a child loves smarties and hates jelly beans. When that child is asked whether he or she wants to have the smarties or the candies, the response may be "smarties" because of the child's preference, and not because of logical inability. The child may not want to have the jelly beans as well as the smarties. But if the choice is jelly beans and candies the child may choose candies in order to get the desired smarties. Thus we would expect more choice of the superordinate class when the less preferred subset is used. This effect would have important methodological implications. To investigate this possibility, the children were pretested for individual subset preference, and half of the children always received their more preferred subset in the inclusion question, and half always received their less preferred subset. This was done for both Siegel's and Piaget's question form, and more or less preferred subset was a between-subject factor in the experimental design.

Two innovations in the procedure for Siegel's inclusion question were instituted during the course of the experiment. In the first case, some of the third grade children seemed reluctant to act greedy in response to Siegel's question, and this observation was supported by post-experiment interviews. Consequently all of the fifth- and seventh-graders were told, "Now there's one rule to this, you've got to tell me what you _really_ want, O.K.?"

The other change in procedure resulted from something one child did: he answered a Siegel inclusion question, and then picked up the items chosen and put them in his cup before the experimenter could do so. It seemed that this deviation in procedure was a behavioural response which could provide useful information, because if a child understands the choice made the verbal response and the response of putting items in the cup should match. From that point on, after their answers to Siegel's questions the children were told to put their choice in their cup.

The second language form examined was based on the consideration that it is difficult to think of real life instances of the Piagetian linguistic form for inclusion. Several times at the dinner table, my children and I discussed this problem in detail, and the best examples we could think of were not very good examples of spoken English:

Please put away your socks or your laundry.

Do you want to have the barbecue chips, or the chips on this dish?

Let's see how well you can count! Are there more little forks, or more forks on the table?

These sentences sound peculiar. We simply do not say things like this to each other in real life situations. However, we do sometimes say things like:

Please put away your socks, or better still, put away all your laundry.

Do you want to have the barbeque chips, or would you like to have all the chips on this dish?

Let's see how well you can count! Are there more of the little forks, or more of all the forks on the table?

In these last sentences the word "all" modifies the superordinate class and makes clear the intended specification of subordinate and superordinate class. If "the As and all the Bs" is the typical form for this kind of expression, then the Piagetian question does not conform to normal language use. Some of the error in inclusion performance may result from misunderstanding the question rather than inability to simultaneously consider different levels of classes.

The hypothesis was that use of the word "all" in inclusion questions would make clear to children that the inclusive comparison was intended, and they would then be able to demonstrate inclusion logic when that logic was available. To test this prediction, half the children received inclusion questions that did not have "all" modifying the superordinate class [Standard condition], and the other half received questions that did have "all" modifying the superordinate class [All condition]. The All form of the Piagetian question was, "Are there more smarties or more of all the candies?", and for Siegel's question the All form was, "Do you want to have the smarties or all the candies?"

A variable of secondary interest in this study was size of subset ratio. Ahr and Youniss (1970) and Isen, Riley, Tucker and Trabasso (1975) found that inclusion performance was better when subset ratio was small, that is, when the larger subset did not have many more items than the smaller subset. To examine the effect of subset ratio, the numerical composition of stimulus sets was varied. Subset sizes were 4:3, 3:2, 4:2 and 5:2, so that the ratios of larger subset A to smaller subset A' were 1.3, 1.5, 2.0, and 2.5 respectively. Because subclass sizes were kept small in this study, the numerosity of the entire superordinate class was not held constant, in contrast to Ahr and Youniss and Isen et al.

To sum up, the general hypothesis in Experiment 1 was that certain linguistic forms would facilitate demonstration of developing inclusion ability, and that this effect would be most apparent at early ages. Specific predictions about inclusion performance were:

1) Performance would be better when "all" modified the superordinate class than when "all" was not used.

2). With younger children there would be more correct response to Siegel's question form, "Do you want to have the As or the Bs?" than to Piaget's "Are there more As or Bs?"

3) Children would be more likely to use subclass comparison on Piaget's standard question than on Siegel's question, or on the "all" modified question.

4) On Siegel's question, children would do better with their less preferred subset, rather than their more preferred subset, used in the inclusion question.

5) As the ratio of A:A' increased, inclusion performance would decline.

Method

Subjects

There were 224 Ss, 16 girls and 16 boys at each of seven age levels. The ages studied were 3, 4, 5, 6, 8, 10, and 12. The preschoolers attended Little People's Day Care Centre and Tapawingo Tribe Day Nursery, and the school children were students at St. Catherine of Siena, St. Theresa, and St. Vincent de Paul schools, all in Hamilton, Ontario.

Boys and girls were ranked separately by age in months, and then were divided into groups of four. For each group, the first child was randomly assigned to one of the four experimental conditions to be described, the second child was randomly assigned to one of the remaining three conditions, the third child to one of the two remaining conditions, and the fourth child was assigned to the last condition. This random-block assignment ensured that average age was about the same in each condition, and there were equal numbers of girls and boys in each experimental group.

Materials

Four kinds of stimulus sets were used. The sets included candies (smarties and jelly beans), marbles (striped and plain), plastic bingo chips (red and green), and pieces of drinking straw (thick and thin). For reasons of safety, pieces of ribbon (velvet and smooth) were substituted for marbles with the preschool children.

Experimental Design

Each S received eight inclusion questions, four of Siegel's and four of Piaget's. Between Ss there were four combinations of set numerosity and content, and two orders of administering Piaget's and Siegel's question form, PPSSSSPP and SSPPPPSS. This resulted in eight combinations, or subject procedures. In each subject procedure, stimulus sets used on the first four inclusion questions were repeated exactly, in content and size, for the second four questions, so that for each specific set a S received one of Siegel's and one of Piaget's questions. An example of a subject procedure is:

trial	form	content	numerosity
1	S	candies	4:2
2	S	marbles	4:3
3	P	chips	5:2
4	P	straws	3:2
5	P	candies	4:2
6	P	marbles	4:3
7	S	chips	5:2
8	S	straws	3:2

Each set of eight subject procedures was repeated in the four cells defined by the two between-subject factors, All and Standard Conditions, and More or Less Preferred subsets. This resulted in 32 subject procedures for each age level. The experimental design is summarized in the following diagram:

	All	Standard
More Preferred		
Less Preferred		

2 orders of Piaget-Siegel questions

X

4 combinations of set size and numerosity =
8 subject procedures in each cell

Procedure

Each $\underline{S}$ came to the experimental session individually. The materials were displayed in a partitioned container, and for each stimulus set the $\underline{S}$ was asked which subset he or she liked better. The $\underline{S}$'s preference determined which subsets were used in the inclusion questions, depending on whether the child was in the More Preferred or the Less Preferred condition.

For each inclusion question the A and A' items were set out in two separate piles. The $\underline{S}$ was asked to name A, A', and B, and then was asked how many items were in each of these classes. For example, "How many smarties are there? How many jelly beans? Candies?" The $\underline{S}$ was corrected if he or she did not indicate that B > A > A', so that counting incompetence would not interfere with the logical question. Within those constraints wrong answers were allowed, since relative and not absolute size was the issue.

The S was next asked if the As were Bs, and if the A's were Bs; for example, "Are the smarties candies? Are the jelly beans candies?" No correction was given to the response.

Finally the class inclusion question was asked. The Piagetian form was, "Are there more As or more (of all the) Bs?" Siegel's form was, "I'm going to give you some of these things, and you get to keep them. Do you want to have the As or (all) the Bs?" In the initial stages of the study, the E put the chosen items in a paper cup after Siegel's question. Later, however, the Ss were told, "OK, put those in your cup," providing a behavioural measure of inclusion performance on Siegel's question. These behavioural response data were obtained at ages 3, 4, 10, and 12 for all Ss; at age 5 for six Ss in the All condition, and five in the Standard condition; and at age 8, for nine Ss in All and eight in Standard condition. No behavioural data were obtained for age 6.

Across trials, the order of A and A' was alternated or counterbalanced in: physical placement of the items; in the questions asking whether the subsets belonged to the superordinate set; and in the inclusion question.

To sum up the procedure for each of the eight inclusion trials, after naming the sets the S was asked how many were in the subsets and the superordinate set, whether the subsets were in the superordinate set, and then the inclusion question was given.

Results and Discussion

A preliminary look at the data suggested that there were no sex differences (Figure 1) and no effects of subset ratio (Figure 2), and

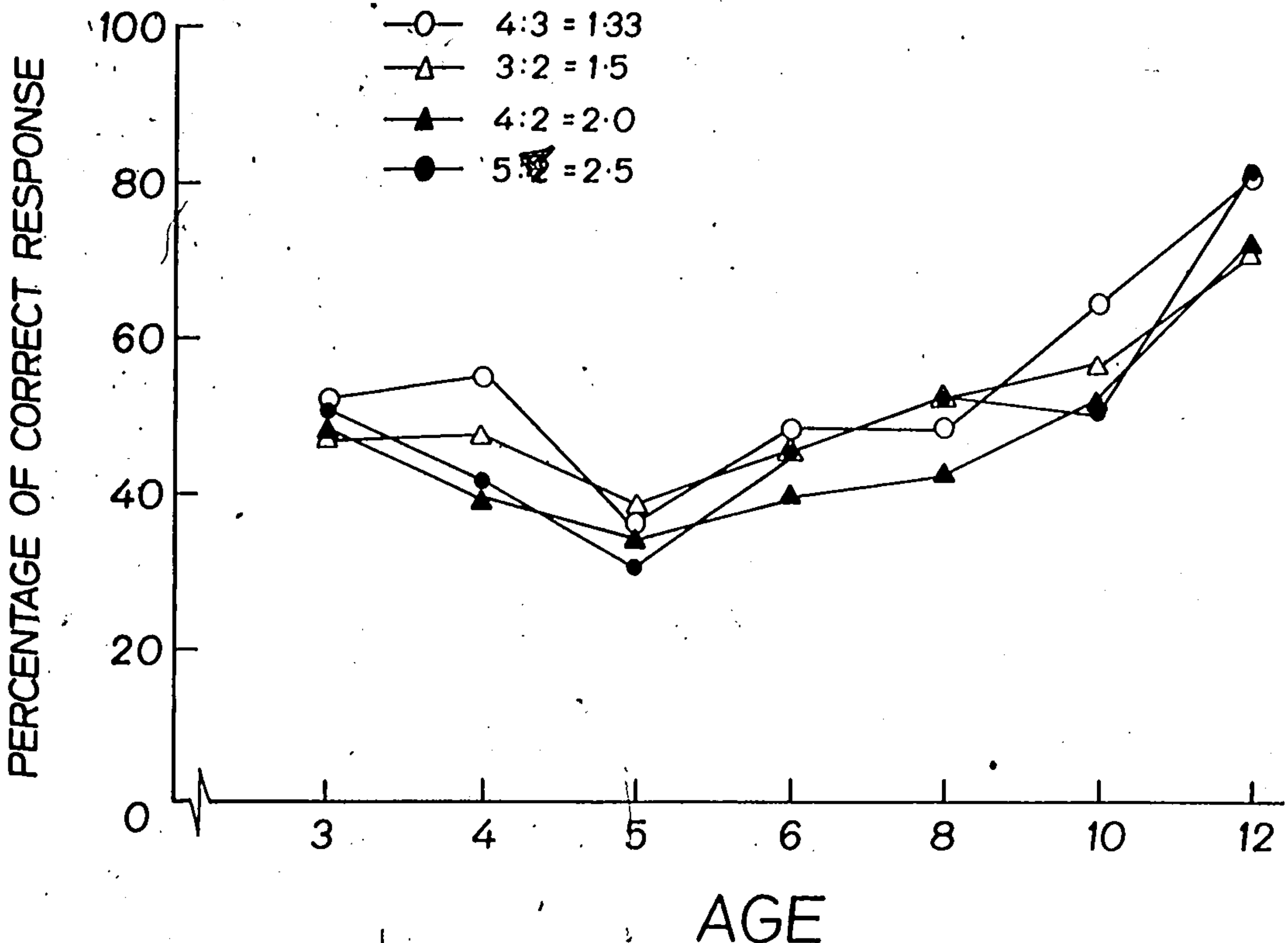
Inclusion Performance of Boys and Girls*

(Experiment 1)

* 16 boys and 16 girls in each age group, 8 questions per S̲.

Figure 1

Effects of Subclass Ratio on Inclusion Performance*



* Each data point represents 32 Ss, 2 responses each

Figure 2

tests were done to determine if this was the case. Since it was possible that boys and girls might have responded differently to the independent variables, particularly the language forms, the number of correct verbal responses was partitioned by sex separately for Siegel, Piaget, All, Standard, More Preferred, and Less Preferred (collapsing across all other factors in each case). Chi-square was calculated for these six comparisons, and also for total correct response of girls and boys across all conditions, for each age level. None of these 49 tests resulted in a significant chi-square (Table 1). There was no reason to expect that effects of subset ratio would be related to any of the independent variables except age, and so the number of correct verbal responses for each of the four ratios was obtained across all conditions. Chi-square was calculated for each age group, and for the different ages combined (Table 2). No effect of subset ratio was found in any of these comparisons. Because there were no sex differences and no effects of subset ratio, these factors were not considered in further analysis.

Different Language Forms

For the major analysis, correct verbal responses were partitioned into Siegel and Piaget question forms, All and Standard conditions, and More and Less Preferred conditions (Table 3). Contingency table analysis was done for each age group, using hypothetical cell frequencies derived from the assumption of no difference for each of the main effects. The results of this analysis are shown in Table 4, with chi-square partitioned into main effects, two way interactions, and three-way interaction.

Table 1

Correct Responses Partitioned by Sex

Age	Sex	Siegel	Piaget	All	Standard	More preferred	Less preferred	Total
3	F	35	27	39	23	35	27	62
	M	33	30	40	23	36	27	63
X_1^2:		.06	.16	.01	0	.01	0	.01
4	F	32	22	33	21	24	30	54
	M	37	25	46	16	30	32	62
X_1^2:		.36	.19	2.14	.68	.67	.07	.55
5	F	32	15	[illegible]	12	21	26	47
	M	31	10	[illegible]	16	16	25	41
X_1^2:		.12	1.00	1.67	.57	.68	.02	.41
6	F	42	22	42	22	27	37	64
	M	34	16	37	13	18	32	50
X_1^2:		.84	.95	.32	2.13	1.80	.36	1.72
8	F	33	33	42	24	27	39	66
	M	34	24	38	20	19	39	58
X_1^2:		.02	1.42	.20	.36	1.39	0	.52
10	F	36	38	41	33	31	43	74
	M	40	28	41	27	34	34	68
X_1^2:		.21	1.52	0	.60	.14	1.05	.25
12	F	47	50	51	46	52	45	97
	M	38	59	50	47	48	49	97
X_1^2:		.95	.74	.01	.01	.16	.17	0

Note: Maximum number of correct responses possible is 64, except for total column where 128 is the maximum.

None of the chi-square values is significant

Table 2

Correct Responses Partitioned by Subset Ratio

Ratio	Age 3	4	5	6	8	10	12	Total
1.3	33	35	23	31	31	41	51	245
1.5	30	30	24	29	33	36	45	227
2.0	31	25	22	25	27	33	46	209
2.5	31	26	19	29	33	32	52	222
χ^2_3:	.15	2.14	.64	.67	.77	1.38	.76	2.95

Note: Maximum number of correct responses possible is 64 for each ratio.

None of the chi-square values is significant.

Table 3

Contingency Tables: Correct Response Partitioned by Question Form, All and Standard Form, and More and Less Preferred Subset

	(Q)	All (A)		Standard (A)		
		More (M)	Less (M)	More (M)	Less (M)	Totals
Age 3	Siegel	25	20	12	11	68
	Piaget	19	15	15	8	57
	Totals	44	35	27	19	125
Age 4	Siegel	24	26	8	11	69
	Piaget	12	17	10	8	47
	Totals	36	43	18	19	116
Age 5	Siegel	17	22	8	16	63
	Piaget	11	10	1	3	25
	Totals	28	32	9	19	88
Age 6	Siegel	27	27	3	19	76
	Piaget	10	15	5	8	38
	Totals	37	42	8	27	114
Age 8	Siegel	18	24	7	18	67
	Piaget	14	24	7	12	57
	Totals	32	48	14	30	124
Age 10	Siegel	21	24	11	20	76
	Piaget	20	17	13	16	66
	Totals	41	41	24	36	142
Age 12	Siegel	20	20	20	25	85
	Piaget	31	30	29	19	109
	Totals	51	50	49	44	194

Note: Maximum cell frequency possible is 32.

Table 4

Chi-Square Partitioned For Matrices in Table 3[a]

Age	3	4	5	6	8	10	12
Siegel/Piaget (Q)	.97	4.17*	16.41***	12.67***	.81	.70	2.97
All/Standard (A)	8.71**	15.21***	11.64***	16.98***	10.45**	3.41	.33
More/Less (M)	2.31	.55	2.23	5.05*	8.26**	1.01	.19
Q x A	.97	3.45	.05	3.51	.03	.25	1.67
Q x M	.20	.04	1.64	.56	.03	1.01	1.32
A x M	.01	.31	.41	1.72	Ø	1.02	.08
Q x A x M	.39	.55	Ø	2.84	.81	Ø	1.01
Total	13.56	24.28	32.36	43.33	20.39	7.41	7.57

a Chi-square values are given in body of table
* $p < .05$
** $p < .01$
*** $p < .001$

There were significant results for all three of the independent variables at certain ages. Children gave more correct response to Siegel's question than to the Piagetian question at ages 4, 5, and 6 (Figure 3). More correct answers were given at ages 6 and 8 when the less preferred subset was used in the inclusion question (Figure 4), and this was true for both Siegel's and Piaget's question. When "all" was used to modify the superordinate class, children gave more right answers than they did to the standard questions at ages 3 through 8 (Figure 5). Thus, all of the main effects were significant at some ages, but chi-square partitioned revealed that there were no interactions between the three factors (see Table 4).

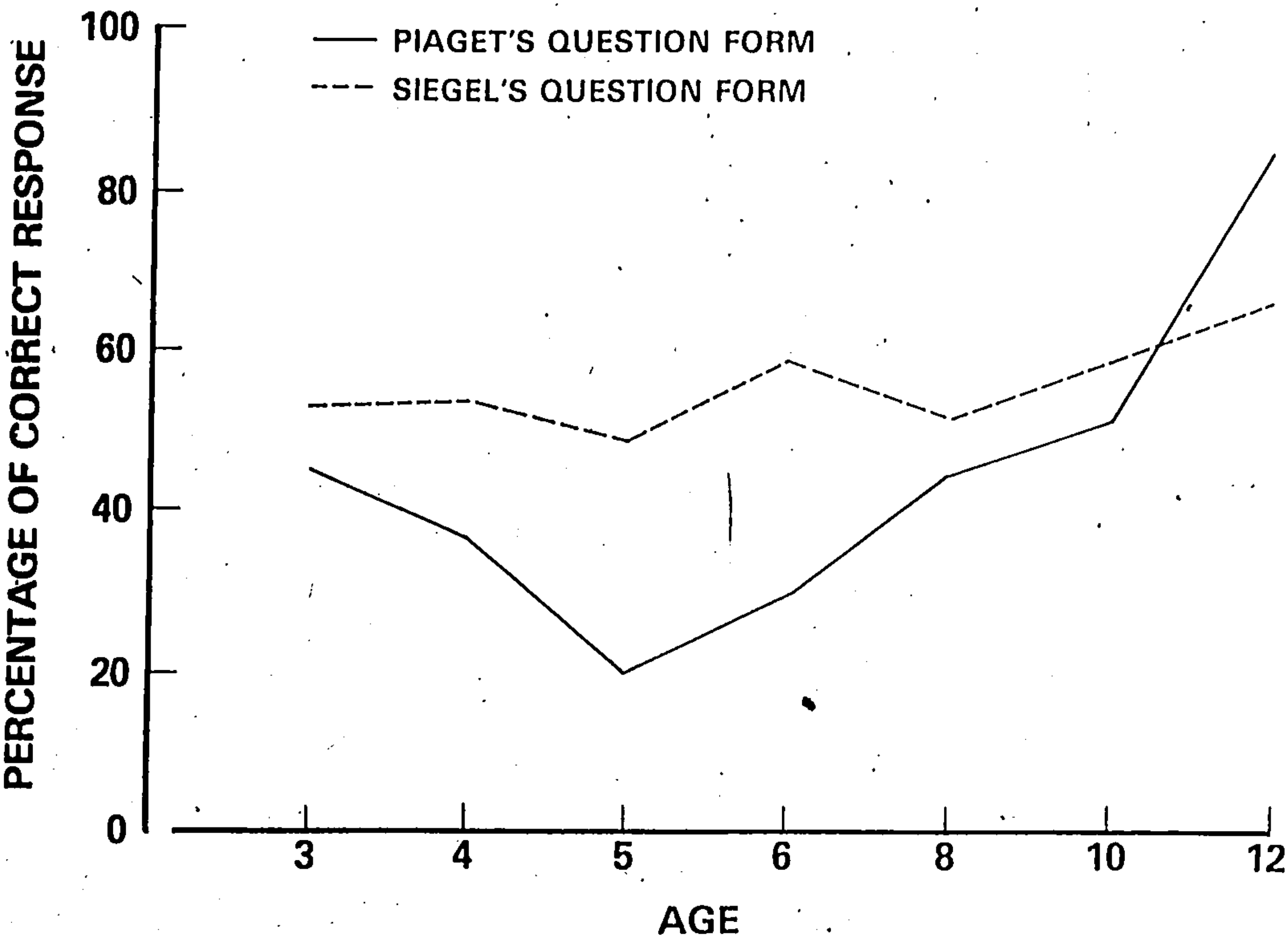
Subset Ratio

In contrast to the results reported by Ahr and Youniss (1970) and Isen et al. (1975), no effects of subset ratio on inclusion performance were found. This may be because total set size was not held constant and because ratios were not varied as widely in this study as in the other two. Here the range of ratios was 4:3, 3:2, 4:2, and 5:2. Isen et al. used ratios of 4:4 and 6:2, and Ahr and Youniss used ratios of 4:4, 5:3, 6:2, 7:1 and 8:0. Perhaps subset ratios affect inclusion performance only with larger sets of items and a sufficiently wide range of ratio variation.

Subset Preference

At ages 6 and 8, the children gave more correct response when their less preferred subset was used in the inclusion question. An unexpected finding was that this effect occurred not only for Siegel's question, when children were asked what they want, but also for Piaget's question when children were asked the quantitative comparison.

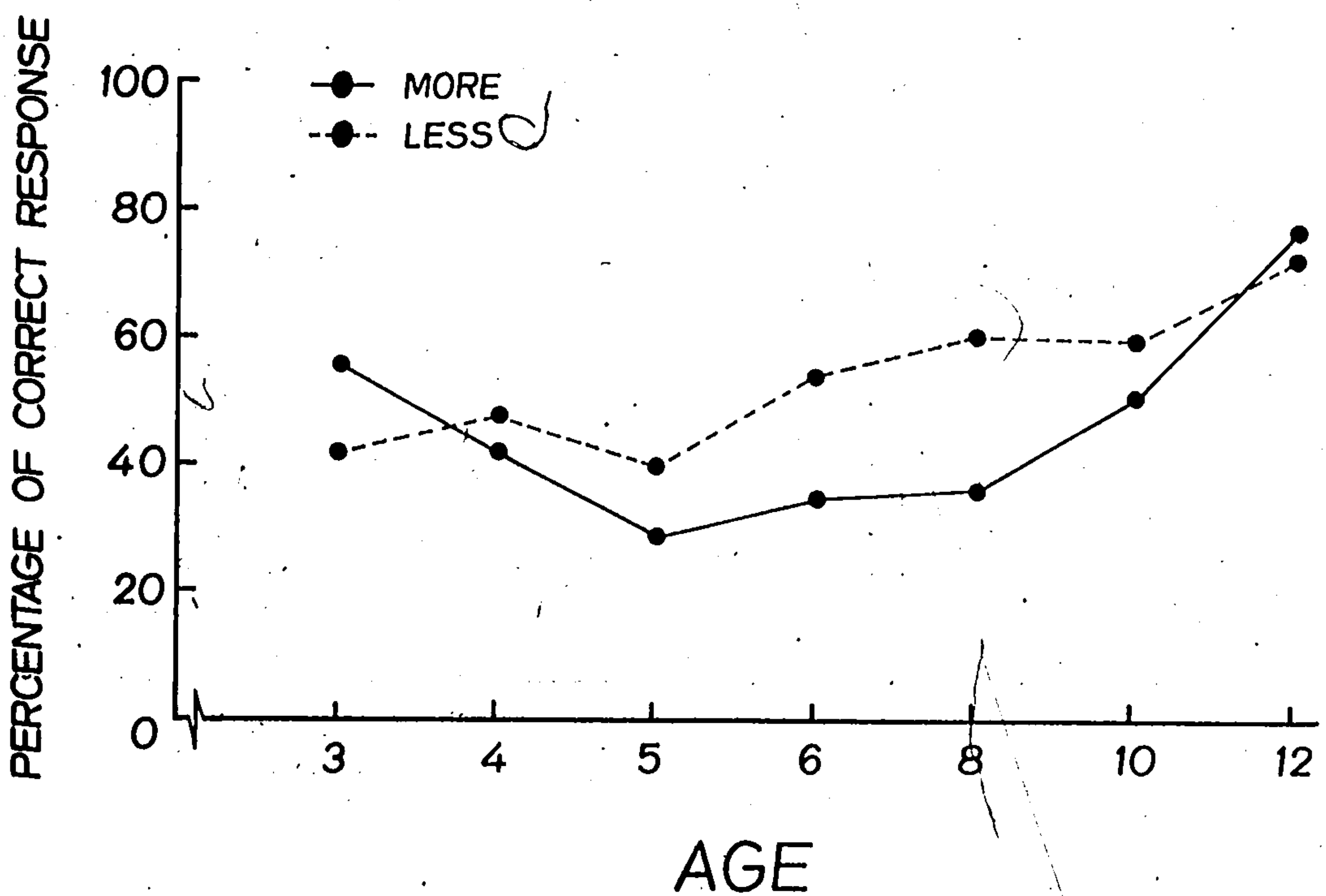
COMPARISON OF SIEGEL'S QUESTION FORM
TO PIAGET'S QUESTION FORM *



* WITHIN-SUBJECT COMPARISON, THIRTY-TWO Ss IN EACH AGE GROUP, FOUR QUESTIONS OF EACH TYPE PER S

Figure 3

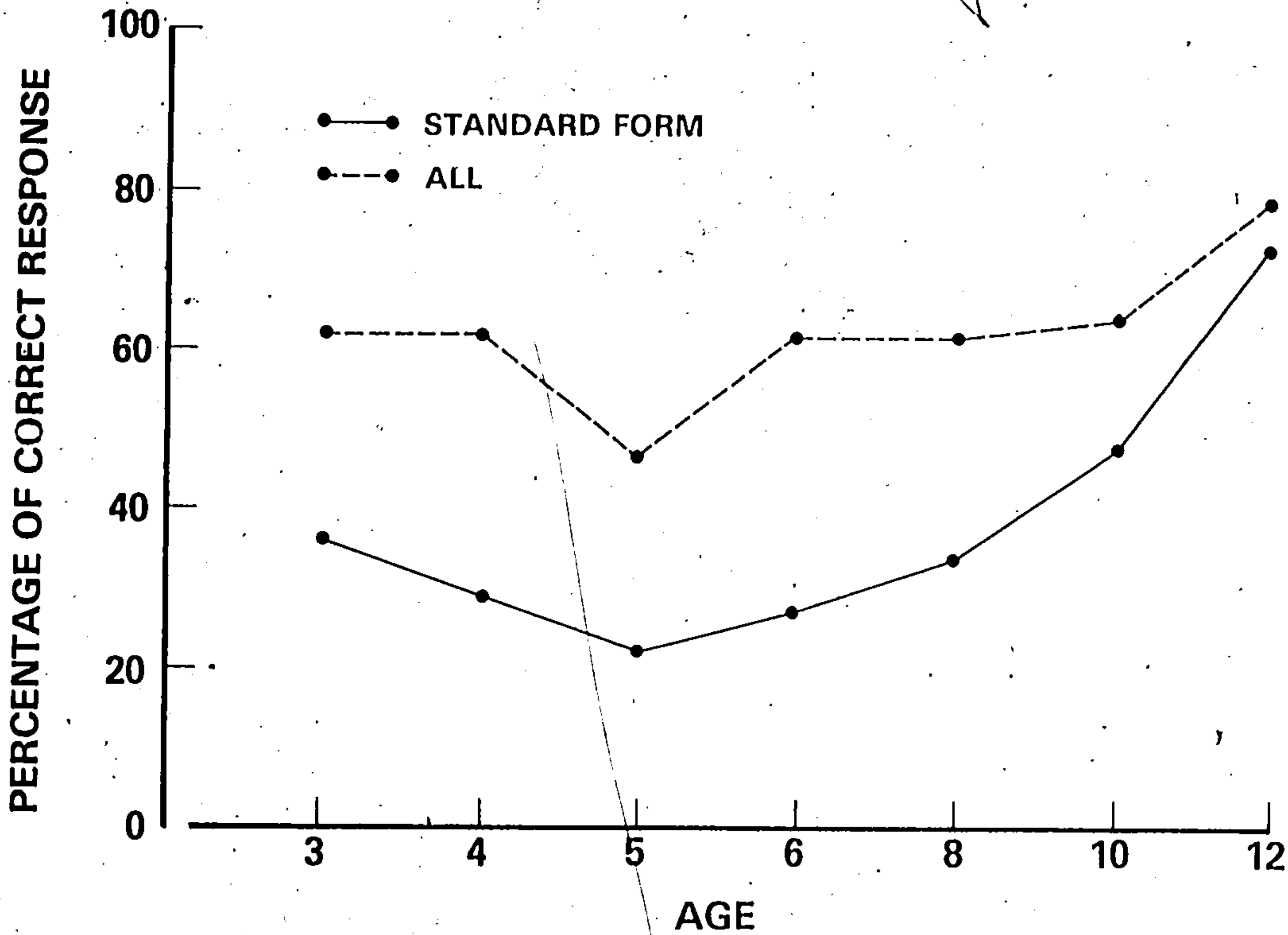
Comparison of More Preferred and Less Preferred Subset*



* Each data point represents 32 Ss, 4 responses each

Figure 4

ALL MODIFYING SUPERORDINATE CLASS COMPARED TO STANDARD UNMODIFIED FORM*



* SIXTEEN _Ss_ IN EACH CONDITION—BY—AGE GROUP, EIGHT QUESTIONS PER _S_

Figure 5

The basis for the effect on Siegel's question appears straightforward: when less preferred subset is one of the choices, the children may choose the superordinate in order to get the items they want most, but when the more preferred subset is one of the choices, then children may be satisfied to have that subset. Choice of the subclass with preferred items, then, would not be logical error but expression of preference. There was no theoretical or practical reason that I was aware of to expect more and less preferred subset to make a difference in response to the Piagetian question, but subset preference did affect performance on this question. Perhaps this is an attentional effect: children might be more likely to attend to and respond with the larger subclass when it is composed of the items they like best.

The effect of more or less preferred subset on inclusion performance is not particularly interesting in itself, but the finding is of some methodological importance. Subjects' preference with regard to the subclasses affects their response on the inclusion question, and so subclass preference should be controlled in inclusion research.

Siegel's and Piaget's Question Forms

Returning to the major hypotheses, Siegel et al. (1977) reported that children of 3 and 4 gave more correct inclusion response to Siegel's question than to Piaget's. This finding is partly replicated in Experiment 1: children 4, 5, and 6 gave more right answers to Siegel's than to Piaget's question form, but there was no difference for the three-year-olds. Percentage of correct response was fairly constant with Siegel's question (Figure 3); for ages 3 through 10, children gave correct answers between 49 and 59% of the time, and this increased to 66% at age 12. It is interesting to note that performance on Piaget's question declined from age 3 to age 5: correct response was 45% for three-year-olds, 37% for four-year-olds, and 20% for five-year-olds. From age 5, Piagetian performance showed consistent improvement, reaching 85% correct response at age 12. These results on Piaget's and Siegel's question forms will be discussed in terms of the strategies children were using.

For children 8 and under, correct response was less than 50% on the Piagetian question and so it appears that some of these children were using a subclass comparison strategy which leads to wrong response. For example at ages 4, 5, and 6, percentage of correct response was 37, 20, and 30% respectively in the Piagetian question, while performance was never below 49% correct response on Siegel's question. Thus, the evidence supports the hypothesis that Piaget's question, in contrast to Siegel's, leads to disjunctive comparison of the two subclasses. Presumably the use of subclass comparison strategy on the Piagetian question is one reason that performance was worse on that question than on Siegel's at 4, 5, and 6.

On Piaget's question, performance was better at ages 3 and 4 than at age 5. This is because some of these young children used a recency strategy: they responded with the last-named class in the inclusion question. At age 3, 9 out of 32 children responded with the last mentioned class on all four Piagetian questions, and at age 4, 5 out of 32 children consistently used this response strategy. By age 5, only two children always chose the last mentioned class on Piagetian questions. Use of a recency strategy by some three- and four-year-olds results in a performance function which is closer to 50% than would otherwise be the case, since the superordinate class was mentioned last in half the questions. The fact that 28% of the three-year-olds used a recency strategy on Piaget's questions explains why no difference was found at this age between the two question forms: performance was close to 50% on both Siegel's and Piaget's questions.

Since performance was between 49 and 59% correct on Siegel's question for ages 3 through 10, one might expect use of the recency strategy for this question as well. However, this was not the case. Only one three-year-old and one four-year-old consistently chose the last-mentioned class on all four Siegel questions, and this would be expected by chance.

There have been previous reports of recency response strategy for young children on logical tasks. On a conservation task, Siegel and Goldstein (1969) found that children two-and-a-half to three-and-a-half chose the last alternative about 90% of the time, and this use of strategy steadily declined until it reached chance levels around four-and-a-half. On class inclusion tasks, Kalil, Youssef, and Lerner (1974) found use of last-mention strategy with somewhat older children. Five- and six-year-olds gave the correct answer 61% of the time when the superordinate was the last class given in the question, and only 33% of the time when the superordinate was the first class given in the question.

At age 12, children gave more right answers to the Piagetian than to Siegel's questions, although this difference was not significant. In post-test interviews some of the older _S_s indicated that they did not choose the superordinate class on Siegel's question because they did not want to act greedy, despite the instructions designed to avoid this contingency. Judging from Ss' behaviour during the experimental sessions it appears that this "modesty" effect began about age 8. Thus with older children, Siegel's question did not reliably measure inclusion logic because of social psychological effects which have no relationship to logical ability.

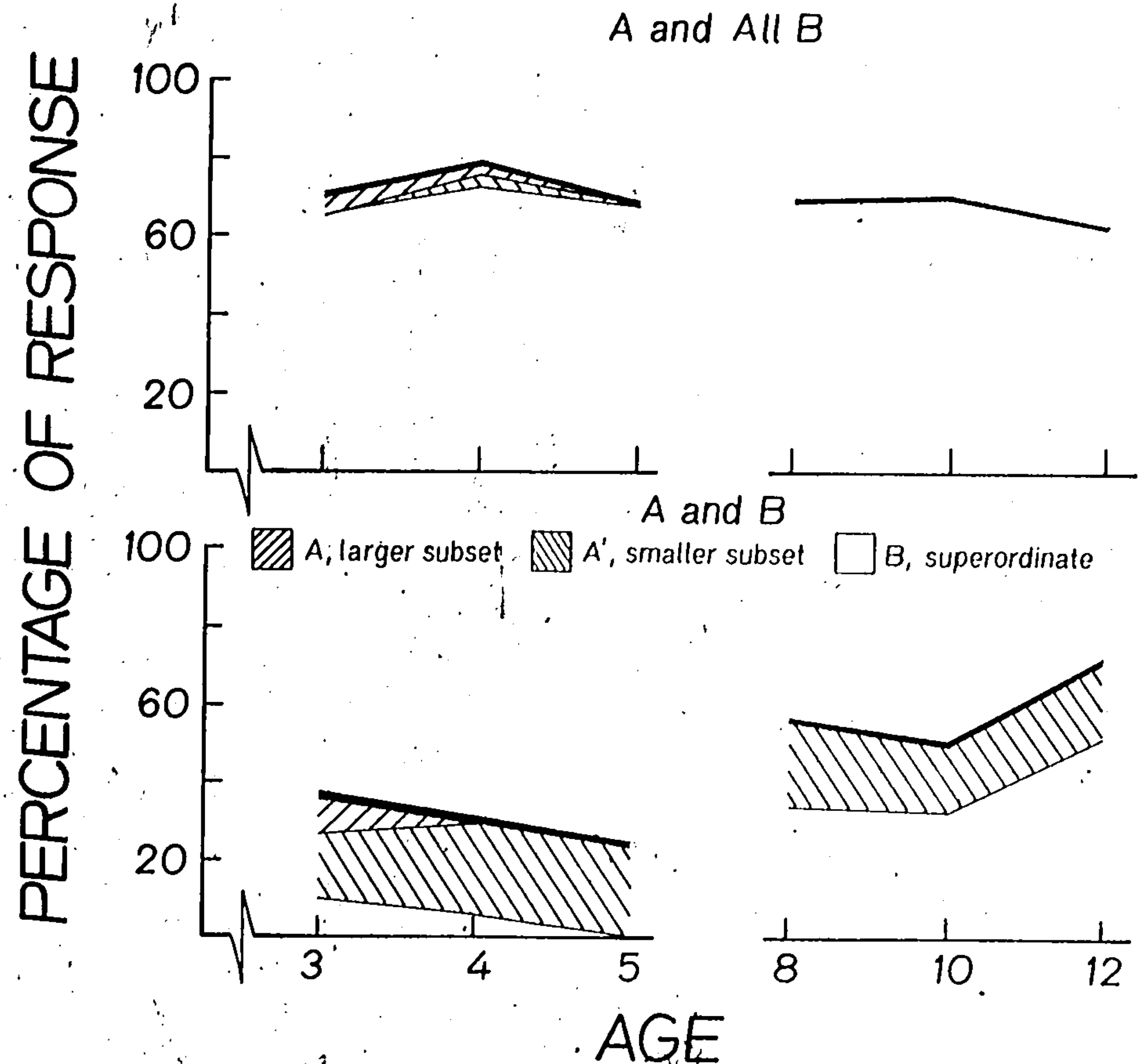
The preceding discussion has been about strategies chidren used which were not related to inclusion logic. On the Piagetian question young children showed some tendency to compare the subclasses to each other, which always leads to the wrong answer, and still younger children sometimes used the strategy of naming the last alternative offered, which leads to correct answers half the time and wrong answers half the time. On Siegel's question, some older children refrained from giving the right answer, naming the superordinate class, because they didn't want to seem greedy.

There is another noninclusion strategy possible which would be represented by correct verbal response on Siegel's question. It could be argued that performance with less preferred subsets in Siegel's question was artificially inflated in the sense that some children were not logically choosing the superordinate class, but were in effect choosing "not the subset offered". That is, offered A and B, where A is the least preferred subset and B is all the items, the children thought "not A" and said "B". If children are using this other-subset strategy, it should be evident on the behavioural measure: after saying B, the children should take the other subset, A', if A' is what they really were choosing. In other words, correct verbal response followed by the erroneous behavioural response of taking the smaller subclass is evidence of the other-subset strategy.

Figure 6 shows verbal and behavioural responses to Siegel's inclusion question in the All and Standard conditions. The heavy line at the top of the graphs shows percentage of correct verbal response, and underneath that line are segments showing the behavioural response which accompanied correct verbal response. In the All condition at age 4, a small percentage of the response demonstrated the other-subset pattern, B verbal response and A' behavioural response. But at the other ages, no child in the All condition gave the correct verbal response and then took the smaller and more preferred subclass. Thus the other subset strategy is not responsible for correct verbal response to Siegel's question in the All condition.

However, the argument that some correct verbal response is based on preference for the smaller subclass appears to have merit for Siegel's questions in the Standard condition. In this case, following correct verbal response many of the three-year-olds and most of the four-year-olds took not the superordinate class they had named but rather the smaller subclass. The behavioural measure was taken for only 5 of the 16 five-year-olds in the Standard condition, but for those children a correct verbal response was always followed by taking the smaller subclass. Many of the chidren aged 8 through 12 did the same thing. Consistent with the reasoning about other-subset strategy, the response pattern of correct verbal response followed by taking only the smaller subclass was most likely to occur when the less preferred

Analysis of Behavioural Response Given Correct Verbal Response on Siegel's Task*



* The top line shows the percentage of correct verbal response, "B," and the accompanying behavioural response is shown by the shaded and unshaded areas under the top line.

Figure 6

subset was used in the inclusion question: across ages, 75% of the cases of this response pattern were with the less preferred subset. Thus, on Siegel's questions in the Standard condition, many of the children were using a strategy of verbally choosing the superordinate class in order to get the items they wanted the most in the smaller subclass.

To sum up the consideration of various strategies used on the inclusion tasks, response to the inclusion question is affected by many factors including comparison of subclasses, order of mention, modesty, and choice of the other, more preferred, subclass. Two of the non-inclusion strategies lead to correct verbal response: the recency strategy used by a few children on the Piagetian question yields correct answers half the time, and the other-subclass strategy, when used on Siegel's question, always resulted in correct answers. To compare the results for the Siegel and Piaget question forms, it seems reasonable to reanalyse the data considering only those correct responses that are not based on either of these two non-inclusion strategies.

Reanalysis of Correct Response on Siegel's and Piaget's Questions

Table 5 shows the correct verbal responses left after eliminating responses based on non-inclusion strategies. On Siegel's question every correct verbal response followed by taking only the smaller subclass was omitted, and correct verbal responses for Piaget's question were omitted for each child who responded with the last mentioned class in all four Piagetian questions. These amended results are tabulated only for three- and four-year-olds because the

Table 5

Correct Responses for Three- and Four-year-olds, Eliminating Responses with Inconsistent Behavioural Responses on Siegel and Order-of-Last-Mention Responses on Piaget

Age 3

	All	Stan.	
Siegel	42	7	49
Piaget	22	17	39
	64	24	88

	x^2		p
Siegel/Piaget (Q)	1.14	>	.05
All/Standard (A)	18.18	<	.001
Q x A	10.23	<	.01
Total	29.55		

Age 4

	All	Stan.	
Siegel	47	4	51
Piaget	21	23	44
	68	27	95

	x^2		p
Q	.52	>	.05
A	17.69	<	.001
Q x A	21.32	<	.001
Total	39.53		

Note: Maximum possible cell frequency is 64.
Order-of-last-mention responses did not occur with Siegel's question. The behavioural measure was not taken for Piaget's question in this study.

behavioural data are not available for all of the older children. This is no problem with regard to the recency strategy, because that strategy was not used after age 4. However, it should be noted that the other-subset strategy in the Standard form of Siegel's question persists through age 12. On Siegel's question when the behavioural measure was taken, all of the correct verbal responses for five-year-olds were followed by inconsistent behavioural response, and this was also true for 39% of the eight-year-olds' correct responses, 34% of the ten-year-olds' correct responses and 25% of the twelve-year-olds correct responses.

Chi-square partitioned was calculated for the three- and four-year-old responses in Table 5. Consistent with the earlier analysis, the children gave more correct inclusion answers in the All than in the Standard condition, and for the three-year-olds there was no difference between Piaget's and Siegel's question. Not consistent with the earlier analysis is the fact that when the non-inclusion responses are eliminated there is no overall difference between Siegel's and Piaget's question for the four-year-olds. Furthermore, for both three- and four-year-olds, there is an interaction between Siegel-Piaget and All-Standard because the children in the All condition did better with Siegel's questions, but in the Standard condition there were fewer correct responses to Siegel's questions.

In conclusion, these findings replicate the report of Siegel, McCabe, Brand and Matthews (1977) that four-year-olds give more correct verbal response to Siegel's inclusion question than they do to the Piagetian question, and the behavioural data indicate that in the All condition the better performance on Siegel's question was based on increased demonstration of inclusion logic. In the Standard condition, however, the children often named the superordinate class on Siegel's question because they were choosing the smaller subclass with more preferred items, not because they were using inclusion logic to choose all of the items. When this kind of response in the Standard condition is eliminated from the data, the children did not perform better on Siegel's question than on Piaget's. It should be pointed out that there were various differences between Experiment 1 and the work of Siegel et al., including a greater variety of materials and presentation of a greater number of questions to the children in Experiment 1, and a change in the wording of the question from "do you want to eat..." to "do you want to have..."

Comparison of the All and Standard Forms

One of the strongest effects found in this study was that children below the age of 10 were more likely to give correct answers to the inclusion question when the word "all" was used in the question. To understand this effect, the behavioural response data for the All and Standard conditions were compared. It should be remembered that in this experiment the behavioural data are only available for Siegel's question, and the behavioural information is incomplete for certain ages.

On Siegel's question, there was not only a high percentage of correct verbal response in the All condition, but the conditional probability of correct behavioural response given correct verbal response was also high (Figure 6). After age 4 there were no errors:

if the children made the correct verbal choice, they put the right items in the cup.

In the Standard condition, Siegel's question, children made behavioural errors following correct verbal response through age 12. From the age of 4, all the errors made consisted of putting the A' subclass in the cup, and, when the behavioural measure was available, five-year-olds did this in every case of correct verbal response. Over 25% of the twelve-year-olds put only the A' items in the cup after naming the superordinate class.

In order to test the differences in behavioural response for the All and Standard conditions post hoc tests were done. The correct verbal response to Siegel's question were partitioned into the accompanying behavioural response, consistent or not consistent with the choice expressed (Table 6). Simple chi-square with the Yates correction was calculated for each age group except the five-year-olds, where the small number of cases required use of the Fisher Exact test. No behavioural response data were available for the six-year-olds. At every age tested, children who made the correct choice were in fact more likely to take all the items in the All condition than in the Standard condition.

A similar post hoc analysis was done of behavioural response when children made the wrong verbal response, that is when they chose the larger subclass (Table 7). At ages 8 through 12, when children chose the larger subclass they always took that subclass, so verbal and behavioural error response was 100% consistent at these ages for both conditions. At 4 and 5, there was no difference between the two

Table 6

All and Standard Conditions: Analysis of Behavioural Response on Siegel's Question When Correct Verbal Response Was Made

	Age 3			Age 4			Age 5		
	All	Stan.		All	Stan.		All	Stan.	
Inconsistent Behavioural Response (A)	3	16	19	3	15	18	0	5	5
Consistent Behavioural Response (B)	42	7	49	47	4	51	17	0	17
	45	23	68	50	19	69	17	5	22
	$X_1^2 = 26.87$			34.31			Fisher exact		
	$p < .001$			$< .001$			$< .0005$		

	Age 8			Age 10			Age 12		
	All	Stan.		All	Stan.		All	Stan.	
Inconsistent Behavioural Response (A)	0	7	7	0	10	10	0	12	12
Consistent Behavioural Response (B)	25	11	36	45	21	66	40	33	73
	25	18	43	45	31	76	40	45	85
	$X_1^2 = 6.76$			14.01			10.32		
	$p < .01$			$< .001$			$< .01$		

Note: Behavioural response data are not available for six-year-olds.

Table 7

All and Standard Conditions: Analysis of Behavioural Response on Siegel's Question When Wrong Verbal Response Was Made.

	Age 3			Age 4			Age 5		
	All	Stan.		All	Stan.		All	Stan.	
Consistent Behavioural Response (A)	8	32	40	13	43	56	7	15	22
Inconsistent Behavioural Response (B)	11	9	20	1	2	3	1	0	1
	19	41	60	14	45	59	8	15	23
	$X_1^2 = 6.02$			.09			Fisher exact		
	$p < .02$			$> .05$			$> .05$		

	Age 8[a]			Age 10[a]			Age 12[a]		
	All	Stan.		All	Stan.		All	Stan.	
Consistent Behavioural Response (A)	11	14	25	19	33	52	24	19	43
Inconsistent Behavioural Response (B)	0	0	0	0	0	0	0	0	0
	11	14	25	19	33	52	24	19	43

[a]Statistical tests not appropriate.

conditions in behavioural response following wrong verbal response, and there were only three instances where the superordinate class was taken. At age 3 however, children in the All condition were more likely to take the superordinate class following incorrect verbal response than were the children in the Standard condition. Thus, when "all" was used in the inclusion question the three-year-olds were likely to take all the items in front of them even when they did not give the right answer to the question.

This finding with the three-year-olds raises the question of exactly what children in the All condition were doing at different ages. Why does the use of "all" lead to more correct verbal response to the inclusion question? Is this effect based on increased demonstration of inclusion logic, or, at some ages, is the effectiveness of "all" based on non-inclusion strategies? This question is addressed in detail in the next chapter.

Chapter 6

Experiment 2: Examination of "All" Effect

Below the age of 10, children are more likely to give correct answers when "all" is used in the inclusion question than in the standard form of the question which does not contain the word "all", and the results of Experiment 1 indicate that this is a strong effect. Why is inclusion performance better with the All form? Is it that children better understand the logical request made, and so are able to demonstrate whatever inclusion competence is available? The behavioural data from Experiment 1 supported this interpretation. By taking all the items, the children showed more understanding of their correct answers in the All condition than in the Standard condition which did not have the word "all".

However, the behavioural data analysis with regard to incorrect answers indicated that for one age group, the three-year-olds, something other than a clear-cut demonstration of inclusion logic was occurring. In the All condition as compared to the Standard condition, the three-year-olds were more likely to give the right answer by naming the superordinate class, but they were also more likely to take the superordinate class following the wrong answer of naming the larger subclass. This inconsistency between correct behavioural responses and incorrect verbal responses is a sign of some confusion on the part of these young children. The inconsistency suggests that even when the three-year-olds did not answer correctly, there was something about the word "all" in the question which predisposed them to take all the items. The question is, what logical basis and what strategies account for these results?

There are three plausible possibilities concerning the effects observed with the All form of the inclusion question:

1. The question with "all" conforms to natural language use and makes clear that the inclusive comparison is requested. Because the children understand the question, they are able to demonstrate inclusion logic when that logic is available.

2. The word "all" is associated with words like "more", "big", "much", that is, it is associated with the concept of magnitude. Because of the connotation of "all", children choose the class modified by "all" regardless of which class is modified.

3. Children interpret "all" to refer to "all the items", which they know is the greatest quantity possible, and so they choose all of the items without comparing subclass and superordinate class. Because of their interpretation, children may sometimes choose all the items even when "all" modifies one of the subclasses.

Only the first possibility, increased question clarity, involves inclusion logic. If children are using an association strategy (2) or an all-the-items strategy (3), they are producing right answers to the question without making an inclusive comparison. It is possible, of course, that the basis for improved verbal response with the All question is different at different ages.

To find out more about the effect of "all" in the inclusion question, children were asked a variety of questions in which "all" was used in different ways.[1] Two of these questions were the same as the All and the Standard questions in Experiment 1, in which the children were asked to compare "A and all the B" and "A and B," respectively. Another inclusion form was the Double All question in which "all" modified both the larger subclass and the superordinate class, so the comparison was for "all the A and all the B." The fourth comparative question was not an inclusion question, but rather asked for comparison of the two subclasses with "all" modifying the smaller subclass, so the form of this All Subclass question was "A and all the A'."

These comparative questions were administered between subjects, resulting in four experimental subject groups. The children in the first three groups, who received inclusion questions (All, Standard, Double All), were also asked a Standard Subclass question comparing "A and A'" in order to provide control data for the All Subclass question.

The three possible explanations for the effect of "all" in the inclusion question lead to different predictions about results on the comparative questions. For example, if children are using the association strategy then they should be more likely to choose the smaller subclass A' in the All Subclass question, "A and all the A'," than in the Standard Subclass question, "A and A'." The smaller subclass A' will be chosen because it is modified by "all." But if children are using an all-the-items strategy, then they should be more

[1]This study was part of a larger study in which the children were first given an unrelated set of inclusion questions, then two seriation tasks, and finally the comparative questions described here. It is possible that performance was affected by the earlier tasks, but presumably any such effect would have been the same for all the comparative questions studied.

likely to choose the superordinate class B on the All Subclass question, even though B is not given as an option in the question. On the All and Double All inclusion questions, all-the-items strategy might result in taking the superordinate class following an incorrect verbal response of naming the larger subclass. In addition, all-the-items strategy should also lead to equivalent performance for the Double All and All conditions: the children should do just as well with "all the A and all the B" as with "A and all the B," since any use of "all" would lead to referencing all of the items.

Method

The materials and procedure in Experiment 2 were the same as in Experiment 1, with the exceptions noted here.

Ss in the study were sixteen boys and sixteen girls from each of four age levels, 4, 5, 6 and 8. The children were students at St. Mark's Co-op Nursery, Dundas, and St. Joseph's School, Hamilton, Ontario. There were not sufficient numbers of three-year-olds available to include this age group in the study. Ss were assigned to four experimental conditions defined by the type of comparative question asked:

Condition	Comparison
Standard	A (larger subclass), B (superordinate class)
All	A, all B
Double All	all A, all B
All Subclass	A (larger subclass), all A' (smaller subclass)

Each S was given four stimulus sets, and thus four comparative questions. As in Experiment 1, half of the Ss received their more preferred subset for the larger subclass in the questions, and half received their less preferred subset. Every S received two Piagetian and two of Siegel's questions. Subset preference and question form (Siegel and Piaget) were varied in Experiment 2 for purposes of methodological control, not as a matter of experimental investigation. Subset ratios were varied in the four questions asked each S to provide variety. There were two combinations of set size and content, and order of question form.

Preliminary questions were the same as in Experiment 1. In the three inclusion question conditions (Standard, All, Double All), after Ss were asked how many items were in each class and whether the subclasses were part of the superordinate, they were asked the Standard Subclass question: "Are there more A or more A'?" Following the preliminaries, all Ss were asked the comparative question determined by their experimental condition. The behavioural measure was taken after verbal response for both the Piagetian and Siegel's questions: the children were told, "Put those in your cup" (Siegel), or "Put those in my cup" (Piaget).

Results and Discussion

As pointed out in the introduction to this study, if children are using an association strategy, choosing the class modified by "all" regardless of which class is modified, then there should be more erroneous choice of A' in the All Subclass question "A and all A'" than in the Standard Subclass comparison "A and A'". This was not the case at age 4, when Ss made more errors (18%) with the Standard Subclass question than with the All Subclass question (9% error). The older children made virtually no errors on either subclass question. Simple chi-square with the Yates correction was calculated for correct and error verbal response to the two subclass questions, and there was no significant difference in response for the two questions at any of the four age levels. Thus, these data do not indicate that children 4 through 8 use an association strategy when "all" is used to modify the smaller subclass in the subclass comparison question. Since performance was very high on the subclass questions, demonstration of an association strategy was unlikely.

If children are taking "all" to mean "all the items" then on the All Subclass question Ss might sometimes say the name of the superordinate class, or they might put that class in the cup, even though the superordinate is not named in the inclusion question. However, on this question there was no S in any group who named the superordinate class, and no S named the smaller subclass modified by "all" and then placed the superordinate class in the cup.

Another possible consequence of an all-the-items strategy is that on the inclusion questions which include the word "all", All and Double All, Ss might show some tendency to put the superordinate in the cup even when they made the wrong verbal response of naming the larger subclass. In Experiment 1, a significant portion of the responses of the three-year-olds in the All condition followed this pattern. In Experiment 2 there was no difference between the All and Double All conditions with regard to behavioural responses following wrong verbal response (Fisher exact tests, all p > .05), so these two conditions were combined for purposes of comparison with the Standard condition (Table 8). The four-, five-, and six-year-olds did make proportionately more inconsistent behavioural responses following incorrect verbal choice in the two "all" conditions than in the Standard condition, but these differences were not significant.

Finally, if children are doing better in the All condition because of all-the-items strategy, then performance should be as good in the Double All as in the All condition. Figure 7 shows the proportion of correct verbal response in the three kinds of inclusion questions with the behavioural responses which accompanied correct answers. Comparing the number of question sets on which both verbal and behavioural response was correct, at age 4 and 5 Ss gave more correct response to the All question than to the Double All question, which indicates they were not using all-the-items strategy 3 (Table 9). There was no difference between All and Double All performance for the six- and eight-year-olds, so it is possible these older children were using an "all the items" strategy. However, since the All question leads to more correct response than the Standard question with the four- and five-year-olds who are not using an all-the-items strategy (Table 10), there is no reason to assume that the older children are using this strategy.

Summary

There was evidence in Experiment 1 that the three-year-olds were using an all-the-items strategy to some extent. But for children older than 3, the results of Experiment 2 indicated that neither an

Table 8

Comparative Questions: Analysis of Behavioural Response When Wrong Verbal Response Was Made

Age 4

	Stan.	Double All and All	
Consistent Behavioural Response (A)	14	23	37
Inconsistent Behavioural Response (B)	1	8	9
	15	31	46

$X_1^2 =$ 1.29

p > .05

Age 5

	Stan.	Double All and All	
Consistent Behavioural Response (A)	19	28	47
Inconsistent Behavioural Response (B)	3	8	11
	22	36	58

.22

> .05

Age 6

	Stan.	Double All and All	
Consistent Behavioural Response (A)	19	19	38
Inconsistent Behavioural Response (B)	0	3	3
	19	22	41

$X_1^2 =$ 1.15

p > .05

Age 8

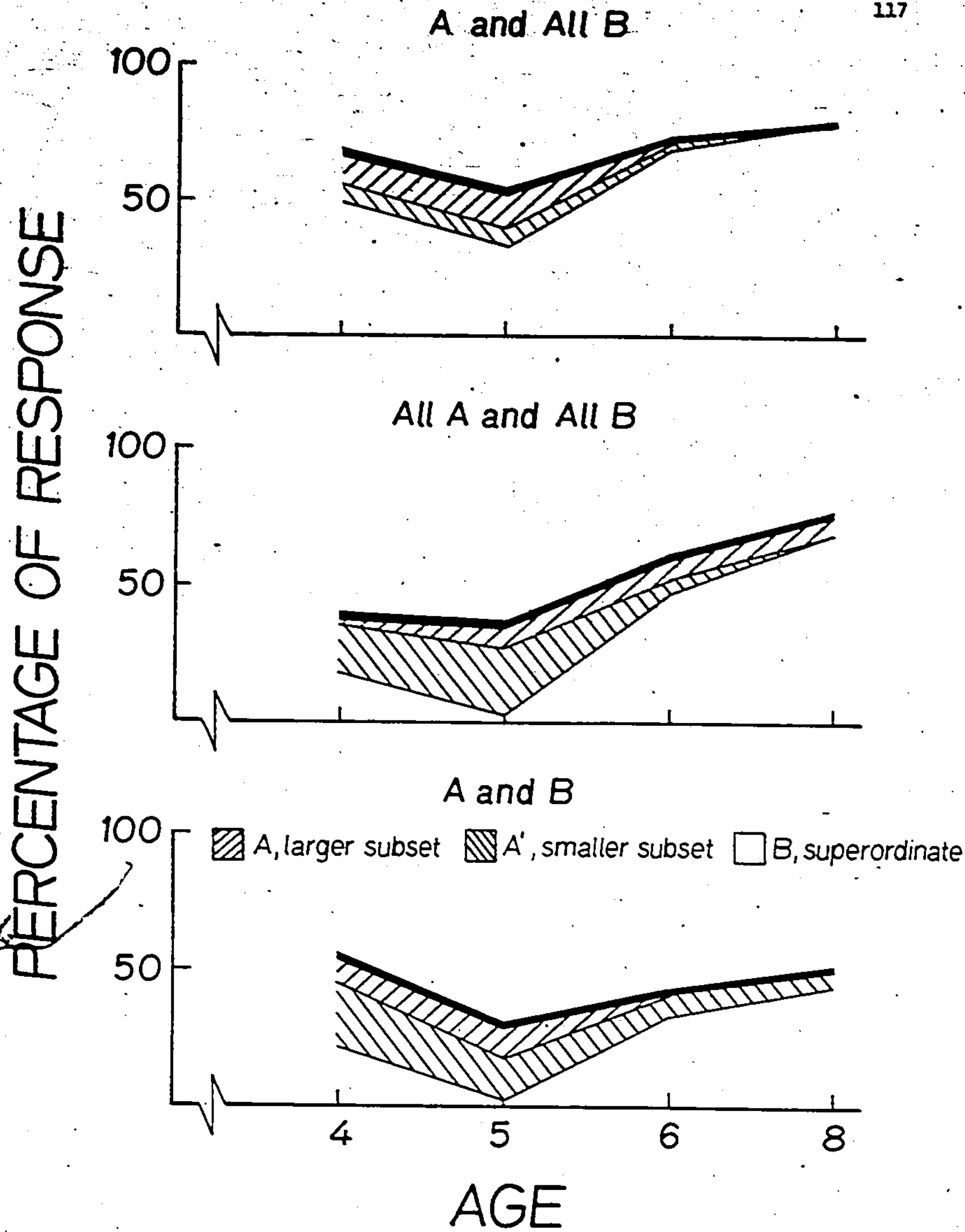
	Stan.	Double All and All	
Consistent Behavioural Response (A)	14	15	29
Inconsistent Behavioural Response (B)	2	0	2
	16	15	31

Fisher Exact test

> .05

Analysis of Behavioural Response

Given Correct Verbal Response*



* The top line shows the percentage of correct verbal response, "B," and the accompanying behavioural response is shown by the shaded and unshaded areas under the top line.

Figure 7

Table 9

Comparative Questions: Correct on Both Verbal and Behavioural Responses to the Double All and All Questions.

	Age			
	4	5	6	8
Double All	6	1	15	22
All	16	11	22	25
Total	22	12	37	47

	Binomial Test		$X_1^2 = 1.32$	$X_1^2 = .19$
p	=.026	=.003	> .05	> .05

Note: Maximum possible correct responses is 32 for each age - comparative question group.

Table 10

Comparative Questions: Correct on Both Verbal and Behavioural Responses to the Standard and All Questions

	Age 4	5	6	8
Standard	7	1	11	14
All	16	11	22	25
Total	23	12	33	39
	Binomial Test		$\chi^2_1 = 3.67$	$\chi^2_1 = 3.10$
	p = .047	= .003	> .05	> .05

Note: Maximum possible correct responses is 32 for each age – comparative question group.

association strategy nor an all-the-items strategy accounts for the facilitative effect of "all" in the inclusion question.

This experiment had certain limitations. There were only eight *S*s in each comparative question condition, and it is unfortunate that three-year-olds were not available for study since Experiment 1 results indicated that age group may be using a non-inclusion strategy. In future work with comparative questions, it would be a good idea to include a separate experimental condition for the Standard Subclass comparison, A and A', and to obtain the behavioural data for that comparison.

This preliminary work with comparative questions, which vary use of the word "all", does not indicate that the effect is a result of "all" being associated with the concept of magnitude, or of "all" being interpreted as "all the items" without inclusive comparison. While further investigation is appropriate, the most plausible assumption at this time is that the All form, "A and all the B," clarifies the meaning of the question. Children are more likely to make the correct inclusion response to the All question because they understand that the inclusion comparison is requested.

Chapter 7

Experiment 3: Reduction of Language

One way to vary language in the assessment of class inclusion is to reduce the language which is used. As mentioned in the introduction, there have been many studies with low verbal assessment of conservation and seriation, but only one study (Braine, 1962) has attempted low verbal assessment of classification. Miller (1976) suggested that perhaps classification is not suited to reduction of language in assessment. However, based on the discussion in Chapter 2, it appears likely that much early classification functioning is not language-bound, nor language-defined, but rather involves non-linguistic representation. Thus, assessment of classification with reduced language demand is appropriate for young children, and it is worth seeking effective means to do this.

Earlier work in low verbal assessment of logical operations reduced the language used in posing the logical demand, or in response requirements, or both. The emphasis was on making sure children understood the question, and were able to perform the response required. In reducing langauge in the question, or in the required response, it may be important to consider critical aspects of the logic in question. In the case of classification the identification of classes is crucial to any classification operation, including class inclusion. If, as suggested in the introductory chapters, non-linguistic category representation is a common and possibly preferred mode of representation for young children, then non-linguistic class representation in assessment is consistent with children's cognitive processes. One way to represent classes non-verbally is by means of visual symbols.

<u>Visual Symbols to Identify Classes: A Negative Request Study</u>

Usually classes are designated by words, which is to say they are named. But sometimes visual symbols identify classes, as in the case of the stylized signs representing woman, man, wheelchair access, deer on the highway. Visual symbols for classes may lead to different cognitive processes than those used when words name classes. Visual symbols would be especially conducive to use of a matching strategy in which one matches instances to the class symbol on the basis of common perceptual attributes. Matching may be a naturally preferred strategy for young children because, as the discussion in Chapter 2 indicates, young children appear to rely on "iconic representation," involving visual imagery (Bruner, 1964), and their sorting behaviour is largely determined by the visual perceptual properties of the stimulus items (Inhelder and Piaget, 1964/1969).

Matching items to a class symbol is probably an effective strategy for identifying positive class instances. However, this strategy should work poorly with "negative requests" for something that is <u>not</u> a member of a specified class, for example, "Give me something that is not a car." The demands of a negative request appear to conflict with the process of matching for common perceptual properties. When words identify classes, on the other hand, the cognitive processing probably involves referencing of class attributes, and these attributes range from concrete, perceptual attributes to relatively abstract attributes such as those based on function or physical composition. As opposed to matching for general visual similarity, cognitive processing based on attributes is more abstract, more flexible in the sense that it is less tied to concrete properties. Negative requests for something that is not a class member are probably more difficult in general than positive requests for a class member; but given the assumptions about the cognitive processes associated with verbal and visual symbol identification of classes, it is further assumed that negative requests are relatively more difficult with visual symbol identification. This assertion is based on the idea that negative requests directly conflict with matching processes, but are less incompatible with processes based on abstraction of class attributes.

If it can be demonstrated that the performance discrepancy between positive and negative requests is greater for children when visual symbols are used than when words identify classes, then support can be adduced for the proposition that different processes are involved in the two cases, and that the process used with visual symbol identification of classes is a perceptual matching process.

Method

Ss were 16 preschool children (mean age 4:1) from St. Mark's Co-op Nursery, Dundas, 16 children in kindergarten (5:10) and 16 in first grade (6:10) at St. Joseph's School, Hamilton, Ontario. From each age level there were equal numbers of girls and boys in the

Standard condition in which words identified classes, and in the Reduced Language condition, with visual symbols identifying classes.

There were six stimulus sets, half with pictures of concrete objects like dolls, cars and dogs (realistic sets) and half with abstract figures defined by shape and color, such as yellow triangles and red crescents (geometric sets). For each of the six stimulus sets, _S_s received three positive and three negative requests, for a total of 36 responses per _S_. Examples of the questions used are:

Standard, positive request: "Give me something that is a car."

Standard, negative request: "Give me something that is not a car."

Reduced Language, positive request: "Give me something that is like this (car symbol)."

Reduced Language, negative request: "Give me something that is not like this (car symbol)."

The number of each kind of item set out was such that the probability of answering correctly by chance was the same for evry question asked the _S_s.

Results and Discussion

There were no differences in the number of correct responses to the realistic and geometric sets (chi-square analysis), nor were there differences in the overall pattern of results for the three age groups. Consequently all _S_s were combined, and analysis concerned only Standard and Reduced Language conditions, positive and negative requests.

Each S received equal numbers of positive and negative requests, and the analysis concerned the relative difficulty of these requests in the two language conditions. Omitting those Ss who gave the same number of correct answers to positive and negative requests, the number of Ss who gave more right answers to positive requests was compared to the number of Ss who gave more right answers to negative requests. In the Standard condition, 13 Ss did better with positive requests and 8 did better with negative requests, and this is not a significant difference, $p = .192$, one-tailed Binomial Test. In the Reduced Language condition the difference is significant: 20 Ss gave more right answers to positive requests and only 4 gave more right answers to negative requests, $p = .001$, one-tailed Binomial Test. Thus, the prediction concerning use of symbols to identify classes was supported. The children did more poorly with negative requests than positive requests when visual symbols were used, but the difficulty with negative requests was not significant when words identified the classes.

This was an exploratory study; one important limitation, in retrospect, is that familiarity of the stimulus materials, or classes, was not varied. I expect that with very familiar materials children will do better with verbal labels on both positive and negative requests, but that with less familiar materials children will do better with visual symbols on positive requests but not on negative requests. In other words, with unfamiliar materials the prediction is that with visual symbols performance on positive requests will be higher and performance on negative requests will be lower than when verbal labels are used to identify classes. This would be a stronger finding than the one in the present study, and would support the assertions that visual symbols predispose to a perceptual matching strategy, and that with younger children that strategy is the most effective for identification of class members. However the finding here, that negative requests are relatively more difficult than positive requests in the visual symbol condition compared to the verbal label condition, does support the assertions that the two modes of class identification lead to different strategies, and that the strategy used with visual symbol identification is based on perceptual matching. Given these results I decided to see whether the mode of class identification, verbal or visual, affected the class inclusion performance of children.

Experiment 3

In Experiment 3, the basic language structure of the standard Piagetian question for testing inclusion was retained, but I identified classes in two ways. In the Reduced Language condition visual symbols designated classes, and in the Standard condition words identified classes. Figure 8 shows a sample stimulus set in the spatial arrangement used in the study. In the Standard condition only the small items comprising the question set were laid out. Children in this condition were asked the standard Piagetian inclusion question: "Are there more blue things or more triangles?" In the Reduced Language condition, symbol cards were used in addition to the stimulus items. The inclusion question in this condition was "Are there more like this (blue symbol) or more like this (triangle symbol)?" Materials and procedure were identical for the two experimental

Stimulus Materials

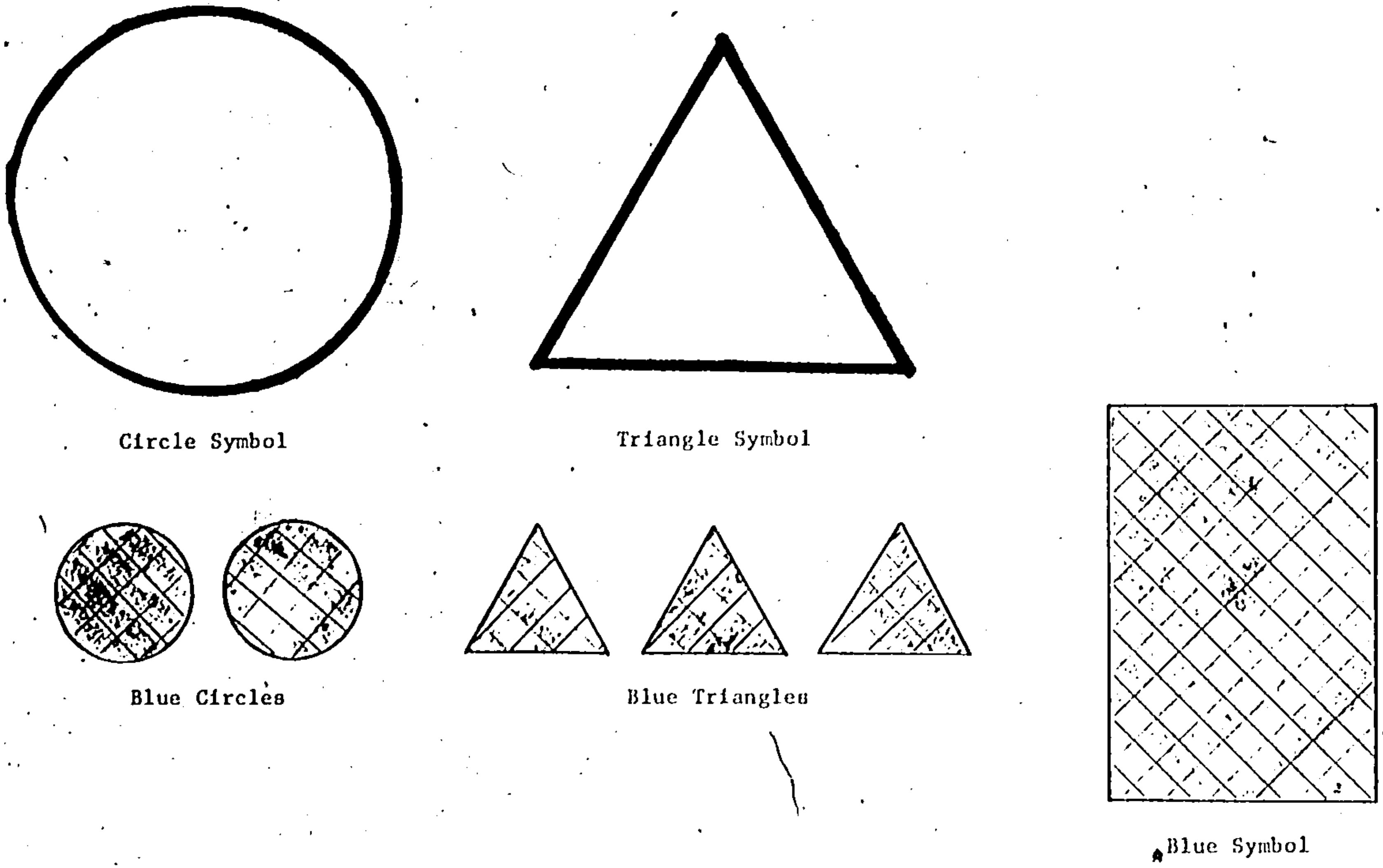


Figure 8

conditions except that symbol cards were used in place of verbal labels in the reduced language variant of the Piagetian question.

The case has already been made that the Piagetian inclusion question may be hard for children to understand. Evidence has been cited (Brooks, 1976; Rosch & Mervis, 1975) that non-linguistic representation is an important part of cognitive functioning, and non-linguistic representation is probably the predominant system in young children (Bruner, 1964). The preceding Negative Instances study indicates that children use different cognitive processes when classes are identified verbally and non-verbally. Perhaps non-verbal identification of classes in the inclusion question will decrease the difficulty of that question for children, allowing demonstration of whatever inclusion logic is available.

Of course, any one attempt to reduce language in the inclusion question may not be successful in the sense of working well for children. If children perform poorly with a reduced language form, or if there is no difference between a standard and a reduced language form, it may be that it is not appropriate to reduce language in classification assessment, or it may be that the approach tried is not a particularly useful approach, and we will not usually be able to distinguish between these two possibilities. But, if children do better on a reduced language inclusion question than on a standard form of the question, then it is clear that the reduced language form does improve children's understanding of the request put to them. Our job in this case is to try to understand the specific effects of the two question forms.

Method

Subjects

The 128 Ss in this study were from four grade levels: preschool (average age 4:7), kindergarten (5:7), first grade (6:7), and third grade (8:7). The preschoolers attended St. Mark's Co-op Nursery, and the older children were students at St. Columba's and St. Cecelia's schools, Hamilton, Ontario. At each age level there were eight girls and eight boys in the Standard condition, and eight girls and eight boys in the Reduced Language condition. Ss were assigned to condition with the random block assignment procedure described in Experiment 1, keeping age equivalent in the two conditions.

Materials

Materials used were figures with the subordinate and superordinate classes defined alternately by shape and color. The six stimulus sets consisted of green and yellow circles, green squares and circles, yellow and blue squares, red triangles and squares, blue and red triangles, and blue circles and triangles. Numerical composition of subclasses was 4-3, 3-1, 3-2, 4-2, 3-2, and 5-2, resulting in subset ratios of 1.3, 3.0, 1.5, 2.0, 1.5, 2.5.

Symbol cards for shape were a white triangle, circle, and square, outlined in black, and symbol cards for color were green, yellow, blue, and red rectangles. The seven symbol cards were prejudged equivalent in perceived "size", and the symbols were larger than the question set items. A sample stimulus set with symbol cards is shown in Figure 8.

Procedure

Stimulus items were set out in a row, one subclass on the left and one subclass on the right, with a space in between. The preliminary procedure was the same as in Experiment 1. Classes were identified, Ss were asked how many items were in each subclass and the superordinate class, and then they were asked whether each of the subclasses was part of the superordinate class. Throughout the task either verbal labels or visual symbols were used depending on each S's experimental condition.

In the Standard condition, classes were identified by pointing to the items specified and saying, "These are triangles, these are circles, these are blue things." In the Reduced Language condition, symbol cards for the subclasses were placed directly above the appropriate subclass items, and the symbol card for the superordinate class was placed to the right of all the stimulus items. The E said, "These are the ones like this (triangle symbol), these are the ones like this (circle symbol), and these are the ones like this (blue symbol)." In both language conditions, the E touched the appropriate items slowly and deliberately while making the identification statements.

For each of the six stimulus sets three comparative questions were asked. In the subclass question, Ss were asked to compare the subclasses to each other (SQ: A r A'?); in the smaller subclass question, they were asked to compare the smaller subclass to the superordinate (SSQ: A' r B?); and in the inclusion question, they were asked to compare the larger subclass to the superordinate (IQ: A r B?).

For the stimulus set shown in Figure 8, the Standard condition questions were:

SQ: "Are there more circles or more triangles?",

SSQ: "Are there more blue things or more circles?",

IQ: "Are there more triangles or more blue things?"

The same questions were asked with visual symbol identification in the Reduced Language condition.

Across stimulus sets the order of the three questions, SQ, SSQ, and IQ, was varied, and the left-right placement of the larger subclass was also varied. The order of the classes within each question was alternated between stimulus sets. The procedure schedule was such that these three variables, order of question type, physical placement, and order of classes in the question, did not covary.

To sum up the procedure, Ss were assigned to either the Standard condition or the Reduced Language condition, and throughout the session classes were referred to either verbally or by visual symbol, depending on the condition. The classes were identified by the E, and then Ss were asked the number of items in each class, and were asked whether the subclasses were part of the superordinate class. Following these preliminaries, three comparative questions were asked, the inclusion question, the smaller subclass question and a question comparing the subclasses to each other.

Results

The correct inclusion responses at each age level are shown in Table 11 partitioned by sex, language condition, and large subset ratio (2.0, 2.5, 3.0) by small subset ratio (1.3, 1.5, 1.7). Contingency table analysis with chi-square partitioned was calculated at each age level (Table 12). No interactions of sex, language condition, and subset ratio occurred. Two of the main effects were significant at one age level only: the girls in first grade made more correct responses than the boys, $p < .05$, (Figure 9), and at age 4, there were more correct responses in the Reduced Language condition than in the Standard condition, $p < .05$. Since the language manipulation was the main point of this experiment, language condition results will be examined in some detail.

Figure 10 shows the percentage of correct response in the Standard and Reduced Language conditions, across ages. By age 8, children in both conditions were making a correct inclusion response more than 90% of the time. Most of the eight-year-olds demonstrated inclusion logic with these very simple, two-dimensional sets. Six-year-olds in both conditions gave the right answer about 50% of the time. Five-year-olds gave more right answers in the Reduced Language condition, but this difference was not significant. The higher proportion of correct response in the Reduced Language condition was significant for the four-year-olds.

The performance function for the Standard condition is similar to the results reported by Piaget, in that the younger children responded correctly less than 50% of the time; their performance was below the chance level. Thus, it appears that many of the younger children in the Standard condition are comparing the subclasses to each other. They say there are more triangles, because there are more triangles than circles. In contrast to Experiment 1, none of the age

Table 11

Correct Response Partitioned by Sex, Language Condition, and Large and Small Subset Ratios

Age 4

	Standard (L)		Reduced Language		
	Small (R)	Large	Small (R)	Large	TOTALS
GIRLS	6	9	7	10	32
—(S)—					
BOYS	6	8	13	17	44
TOTALS	12	17	20	27	76

Age 5

	Standard (L)		Reduced Language		
	Small (R)	Large	Small (R)	Large	TOTALS
GIRLS	4	14	12	13	43
—(S)—					
BOYS	12	7	12	12	43
TOTALS	16	21	24	25	86

Age 6

	Standard (L)		Reduced Language		
	Small (R)	Large	Small (R)	Large	TOTALS
GIRLS	17	15	17	17	66
—(S)—					
BOYS	11	9	10	13	43
TOTALS	28	24	27	30	109

Age 8

	Standard (L)		Reduced Language		
	Small (R)	Large	Small (R)	Large	TOTALS
GIRLS	23	24	23	22	92
—(S)—					
BOYS	23	21	20	22	86
TOTALS	46	45	43	44	178

Note: Maximum possible cell frequency is 24.

Table 12

Chi-Square Partitioned for Matrices in Table 11[a]

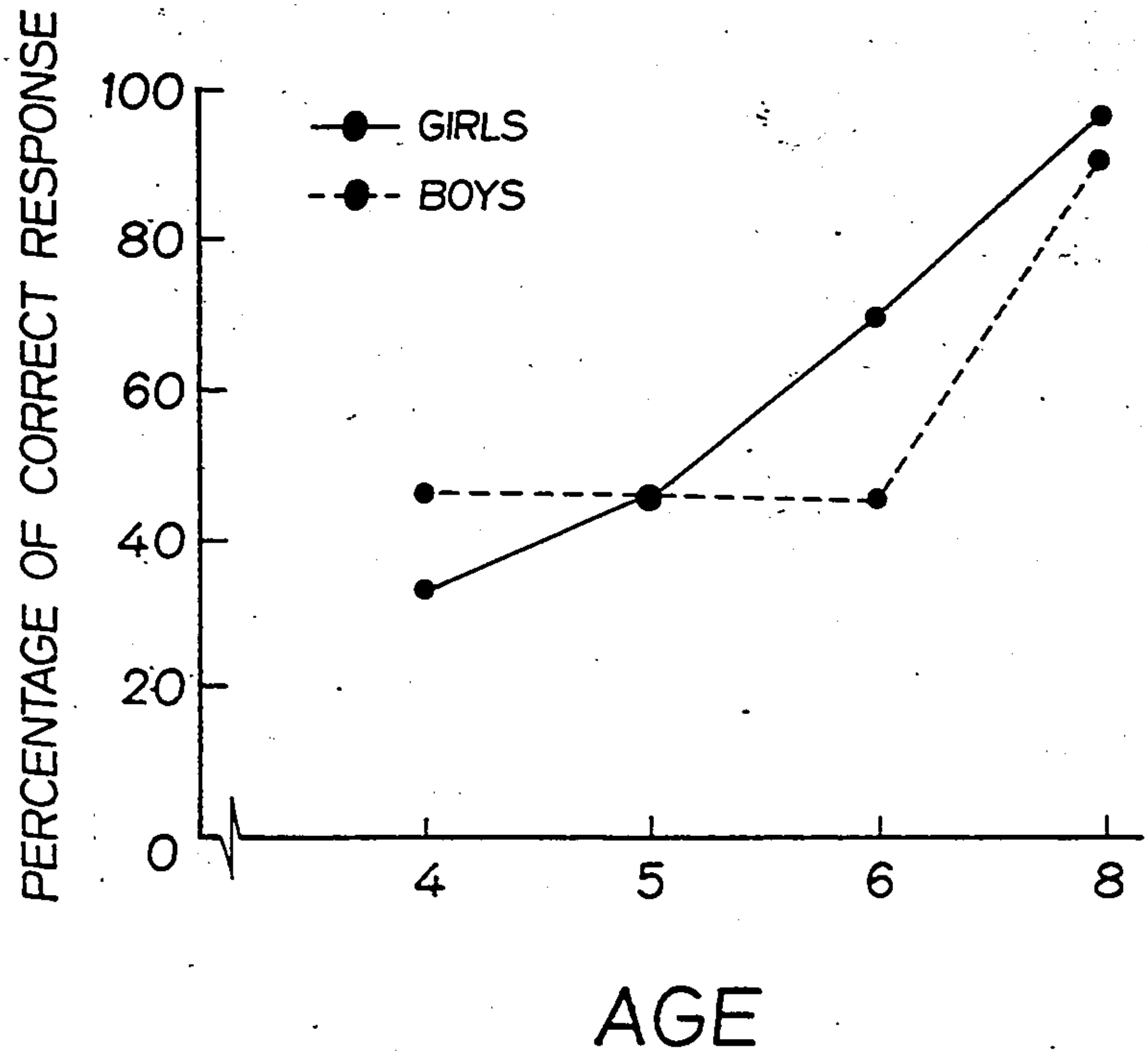
	Age 4	5	6	8
Sex (S)	1.90	0	4.85*	.20
Language condition (L)	4.26*	1.67	.23	.09
Ratio (R)	1.90	.42	.01	Ø
S x L	2.58	.05	.01	Ø
S x R	Ø	2.98	.08	Ø
L x R	.05	.19	.45	.02
S x L x R	.05	2.28	.08	.20
Total	10.74	7.58	5.72	.52

[a]Chi-square values are given in body of table.

*$p < .05$

Inclusion Performance of Boys and Girls*

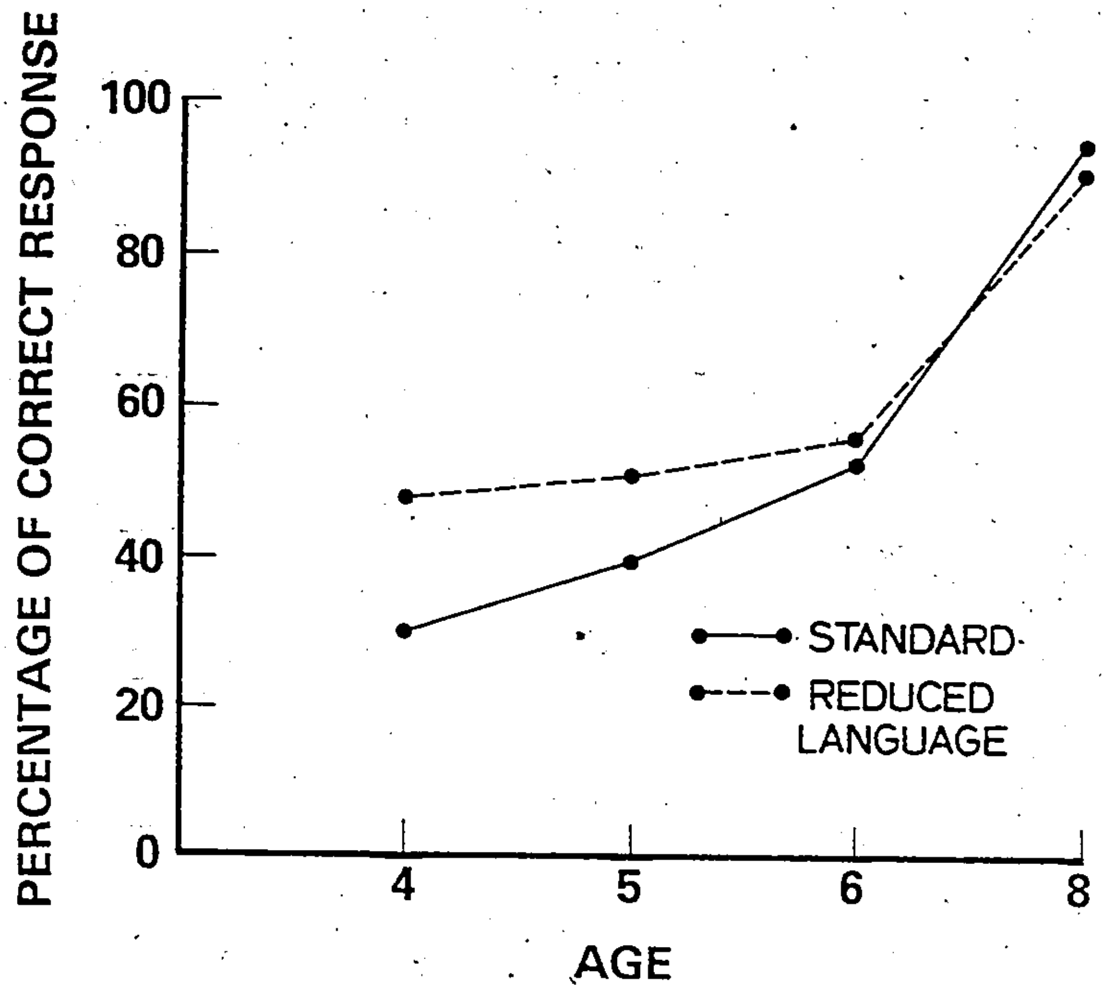
(Experiment 3)



* 16 girls and 16 boys in each age group, 6 responses per S

Figure 9

Comparison of Standard and Reduced Language Conditions*



* SIXTEEN Ss IN EACH LANGUAGE CONDITION—BY—AGE GROUP, SIX SETS PER S

Figure 10

groups in this study relied on an order-of-last-mention strategy. For example at age 4 only one S in each of the language conditions chose the last mentioned class on all six inclusion questions. Of course, the three-year-olds were most likely to use this strategy in Experiment 1, and this age group was not available to participate in Experiment 3.

In the Reduced Language condition the younger children were right about half the time; correct response was 48 to 56% at ages four through six. It could be that all of the children in this condition are guessing, or it could be that half of the children are systematically making the subclass comparison error while the other half are systematically responding on the basis of inclusion logic. Or the basis for performance may be some combination of these two possibilities.

Similarly, in the Standard condition we cannot know at this point the exact basis of performance. Because correct response is less than 50% at ages 4 and 5, it appears that these children are using subclass comparison, which leads to error. But there could be varying proportions of children using inclusion logic, comparison of subclasses, and guessing.

For both conditions, it would be useful to have a better idea of how much response is based on subclass comparison, inclusion logic, and guessing. The mathematical model discussed in the next section provides this analysis.

Chapter 8

Mathematical Model Analysis of Components of Response[1]

In Experiment 3 the children were asked the inclusion question, comparing the larger subclass to the superordinate class, and they were also asked a question comparing the smaller subclass to the superordinate. When children are using inclusion logic they will, of course, give the right answer to both kinds of question. When children guess they will be right half the time and wrong half the time. But when children are comparing the subclasses to each other they will consistently give opposing answers to the two kinds of questions: on the inclusion question (more A or more B) they will give the wrong answer "A" because A is the larger subclass, but on the smaller subclass question (more A' or more B), because A' is the smaller subclass they will respond with the spoken alternative to A', "B," and this is the correct answer. It seemed to me that the consequences of subclass comparison strategy, wrong answers on inclusion and right answers on the smaller subclass question, provided a key to understanding the basis of inclusion performance. Pursuing this line of reasoning, I developed a mathematical model which estimates the underlying components of inclusion response to inclusion questions (Figure 11).

---

[1]Daphne Maurer joined me in discussions leading to the development of this model, and Steve Link gave encouragement and consultation on the testing of the model.

THEORETICAL COMPONENTS OF PERFORMANCE

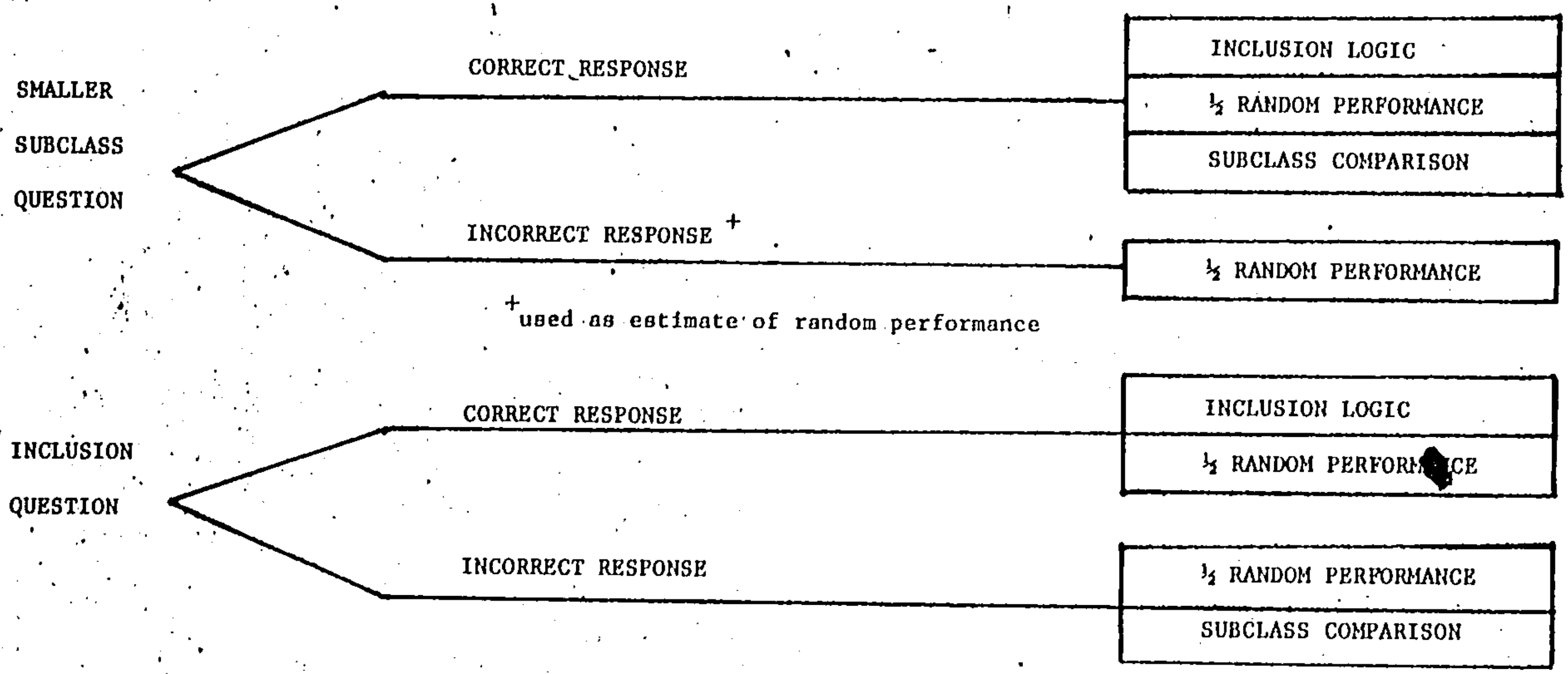


Figure 11

Analysis of Components of Inclusion Performance

Let A and A' be two classes, such that:

$A \cup A' = B$, and,

$A > A'$.

Let SSQ: smaller subclass question, "Are there more A' or more B?"

IQ: inclusion question, "Are there more A or more B?"

Let C: correct response,

E: error response.

Let IL: response based on inclusion logic, comparing specified subclass to superordinate class,

SC: response based on subclass comparison, comparing the two subclasses to each other,

G: response based on guessing, chance performance.

Let p: probability that response is based on IL,

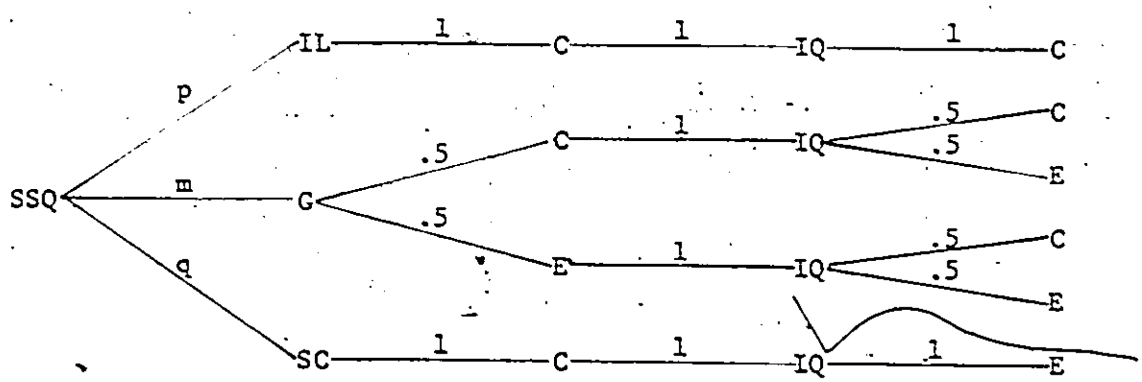
q: probability that response is based on SC,

m = (1-p-q): probability that response is based on G.

Inclusion logic will lead to correct response on both types of questions, but subclass comparison will lead to correct response on the smaller subclass question and to incorrect response on the inclusion question. It is assumed that guessing will lead to correct response half the time and to error half the time, on both types of questions.

There are two further assumptions in this model. The first is that inclusion logic, subclass comparison, and guessing are the only reasonable bases for response to the two comparative questions. The other assumption is that for any one stimulus set a child's performance on the smaller subclass question and the inclusion question is based on the same underlying component, either inclusion logic, or subclass comparison, or guessing, although the basis of performance will not necessarily remain the same for the child on all six stimulus sets. Given the assumptions of this model, the following tree diagram shows the probabilities associated with the outcomes which are possible for the inclusion question and the smaller subclass question in one stimulus set.

<u>Possible Outcomes for Inclusion Question and Smaller Subclass Question in One Stimulus Set</u>



With probability p the child uses inclusion logic, which always leads to correct answers on the smaller subclass question and on the inclusion question. With probability m the child guesses, or performs in a random manner, and this leads to correct answers half the time and wrong answers half the time on both questions. With probability q the child compares subclasses to each other which always leads to correct answers on the smaller subclass question and always leads to error on the inclusion question. Summing across the branches of the diagram, the probabilities for correct and error response on the smaller subclass question and the inclusion question are:

$$pr(C_{SSQ}) = p + q + m/2,$$

$$pr(E_{SSQ}) = m/2,$$

$$pr(C_{IQ}) = p + m/2,$$

$$pr(E_{IQ}) = q + m/2.$$

Incorrect response on the smaller subclass question can be used as an estimate of guessing, since error on this question represents half of the guessing-based response. Given this estimate of guessing, performance based on inclusion logic and subclass comparison can also be estimated:

$$G = 2E_{SSQ},$$

$$IL = C_{IQ} - E_{SSQ},$$

$$SC = E_{IQ} - E_{SSQ}.$$

Validity of the Inclusion Model

In order to consider the validity of the inclusion model, the probabilities for correct and error outcomes shown in the tree diagram are expressed in matrix form:

		Inclusion Question C	Inclusion Question E	
Smaller Subclass Question	C	[a] p+m/4	[b] q+m/4	p+q+m/2
	E	[c] m/4	[d] m/4	m/2
		p+m/2	q+m/2	1

The marginal values in this matrix are the proportion of overall correct and error response on the smaller subclass question and on the inclusion question, and these are the values used to estimate the components of performance in the inclusion model. The values in the cells of the matrix represent the probabilites for the various joint outcomes, such as correct on both questions, or correct on the smaller subclass question and wrong on the inclusion question. The probability values in the matrix are multiplied times the total number of responses to obtain expected cell frequencies.

The obtained cell frequencies are calculated by counting the number of question sets which satisfy the four conditions, that is the number of sets on which Ss are right on both questions (cell a), right on the smaller subclass question and wrong on the inclusion question (cell b), and so on. To examine the validity of the inclusion model, the cell values predicted by the model can be compared to the cell values obtained. This check for validity is based on the discrepancy between obtained frequencies and frequencies predicted by the model.

In the inclusion model, a basic assumption is that an individual child's performance on both questions in a stimulus set will be based on the same underlying component. An alternate assumption is that there is no relationship between a child's response on the smaller subclass question and on the inclusion question. (Note that the probabilites of using inclusion logic, subclass comparison, and guessing, p, q, and m, remain the same with this assumption of no relationship between response on the two questions in a stimulus set.) With this assumption of independence, predicted cell values are products of the marginal values. For example,

cell a = $\frac{(a+b)(a+c)}{a+b+c+d}$, that is, the number correct on both questions,

$$C_{SSQ} \cup C_{IQ} = \frac{C_{SSQ} \times C_{IQ}}{\text{total responses}}.$$

Empirically obtained cell values can be compared to the values predicted by independence, just as comparison can be made to the values predicted by the inclusion model. Calculation of the theoretical cell values of the independence model allows examination of the manner in which these values deviate from the theoretical cell values predicted by the inclusion model.

cell a

$$p + q + m = 1, \text{ therefore } p + q + m/2 = 1 - m/2$$

$$(1 - m/2)(p + m/2)$$

$$= p + m/2 - pm/2 - (m/2)^2$$

$$= p + m/2(1 - p - \frac{1 - p - q}{2})$$

$$= p + m/4(2 - 2p - 1 + p + q)$$

$$= p + m/4(1 - p + q)$$

cell b

$$(1 - m/2)(q + m/2)$$

$$= q + m/2 - qm/2 - (m/2)^2$$

$$= q + m/2(1 - q - \frac{1 - p - q}{2})$$

$$= q + m/4(2 - 2q - 1 + p + q)$$

$$= q + m/4(1 + p - q)$$

cell c

$$m/2(p + m/2)$$

$$= m/4(2p + 1 - p - q)$$

$$= m/4(1 + p - q)$$

cell d

$$m/2(q + m/2)$$

$$= m/4(2q + 1 - p - q)$$

$$= m/4(1 - p + q)$$

For each cell, the value predicted by independence differs from that predicted by the inclusion model by a factor involving p and q which is multiplied times the term m/4. For cells (a) and (d) the factor is $(1 - p + q)$, and for these cells the results of the two models will be the same when $p = q$, the prediction of the independence model will be smaller than that of the inclusion model when $p > q$, and larger when $q > p$. For cells (b) and (c), the prediction of the independence model will be larger than that of the inclusion model when $p > q$, and smaller when $q > p$. In other words, at age levels where performance is about equally likely to be based on inclusion logic and on subclass comparison there will be little difference in the predictions of the two models. But at younger ages, when subclass comparison is more likely than inclusion logic, the independence model predicts more question sets with both questions correct and with both questions wrong than does the inclusion model. Later, when inclusion logic is more likely than subclass comparison, the independence model predicts fewer instances with both questions correct and with both questions wrong.

Given these considerations of how the inclusion model and the independence model differ, if a sufficient age range is available it is possible to consider which model is more appropriate. If the data fit better the predictions of the inclusion model, across ages, support is offered for using that model to estimate the components of class inclusion performance. A further consideration, of course, is how well the data fit the inclusion model.

Table 13 shows the frequencies obtained, the frequencies predicted by the independence model, and the frequencies predicted by the inclusion model, for each age-by-language condition group. The differences in predictions for the two models conform to the differences expected. The independence predictions for cells (a) and (d) are larger than inclusion model predictions at the earlier ages when there is more subclass comparison than inclusion logic, age 4 in the Reduced Language condition and ages 4 and 5 in the Standard condition. At subsequent ages, when there is more inclusion logic than subclass comparison, the predictions of the independence model are smaller than those of the inclusion model for cells (a) and (d). The difference between the predictions of the two models is least at the ages when use of inclusion logic approximately equals use of subclass comparison, ages 4 and 5 in the Reduced Language condition and age 6 in the Standard condition. The differences found between the two models conform to those expected.

Table 13

Frequency Matrices: Predicted and Obtained Values

Form of matrices:

		Inclusion Question C	Inclusion Question E
Smaller Subclass Question	C	Frequency values: predicted by independence obtained predicted by inclusion	
	E		

age	Standard Condition			Reduced Language Condition		
4	22.35	51.65		29.86	31.14	
	20	54	74	33	28	61
	18	56		29.50	31.50	
	6.65	15.35		17.14	17.86	
	9	13	22	14	21	35
	11	11		17.50	17.50	
	29	67	96	47	49	96
5	28.52	45.48		36.24	34.76	
	26	48	74	40	31	71
	26	48		36.50	34.50	
	8.48	13.52		12.76	12.24	
	11	11	22	9	16	25
	11	11		12.50	12.50	
	37	59	96	49	47	96
6	51.46	43.54		42.16	28.84	
	52	43	95	44	27	71
	51.50	43.50		44.50	26.50	
	.54	.46		14.84	10.16	
	0	1	1	13	12	25
	.50	.50		12.50	12.50	
	52	44	96	57	39	96
8	89.10	4.90		81.56	8.44	
	90	4	94	84	6	90
	90	4		84	6	
	1.90	.10		5.44	.56	
	1	1	2	3	3	6
	1	1		3	3	
	91	5	96	87	9	96

In the age-by-language condition matrices, the frequencies predicted by the inclusion model are closer to the obtained values than those predicted by independence in seven out of eight instances. Because there is little difference in the predictions of the two models at certain ages, and because expected cell frequencies are very small in some cases, the overall goodness-of-fit was tested across age levels. Table 14 shows the frequencies obtained, and those predicted by independence and by the inclusion model, for all age and condition groups combined. Chi-square was calculated for both models. The data do not differ significantly from the predictions of the inclusion model ($X^2_1 = 2.88$, $p > .05$), but the data do differ from the predictions of independence ($X^2_1 = 15.56$, $p < .001$). The independence model does not fit the data as well as the inclusion model, and the fit of the inclusion model to the data is satisfactory. Thus, support is offered for the validity of the inclusion model, and it is appropriate to use the inclusion model to estimate the components of inclusion performance.

<u>Estimation of Components of Performance in Experiment 3</u>

When language was reduced in the Piagetian inclusion question by using a visual symbol to identify the classes, two differences in performance were found. At age 6, girls gave more correct answers than boys in both the Standard and Reduced Language conditions, and at age 4 the children gave more right answers in the Reduced Language than in the Standard condition. Since contingency table analysis with chi-square partitioned was calculated for four age levels, one might argue that two significant findings out of the many comparisons could occur

Table 14

Test of Fit: All Age and Condition Groups Combined

		Inclusion Question C	Inclusion Question E	
Smaller Subclass Question	C	368.32 389 380	261.68 241 250	630
	E	80.68 60 69	57.32 78 69	138
		449	319	768

Independence Model: $X_1^2 = 15.56$, $p < .001$

Model I: $X_1^2 = 2.88$, $p > .05$

Note: Form of matrix is the same as in Table 13: frequency values predicted by the independence model are on the top line in each cell, obtained frequencies are on the middle line, frequency values predicted by the inclusion model are on the bottom line.

by chance, and that these are not valid differences. To consider what these differences mean, performance was partitioned into estimated components by means of the mathematical model.

Sex Differences at Age 6

Girls did not give more correct answers than boys at each of the age levels (see Figure 9), and so the superiority of female performance at age 6 was not part of a consistent pattern. The data were partitioned by sex and language condition, and the components of performance were estimated (Table 15, Figure 12). Because the data were divided by sex of subjects, there were only 8 children, or 48 responses, in each group analysed. A negative value was obtained for the estimation of subclass comparison strategy by boys in the Reduced Language condition at age 4. This is because this group of Ss gave more correct answers to the inclusion question than to the smaller subclass question, an outcome which should not occur if children are using either of the systematic strategies, inclusion logic or subclass comparison. The superiority of the inclusion question results over the smaller subclass question results is probably a chance occurrance, an interpretation which is supported by the estimate that 92% of the performance for this group is based on guessing.

In the Standard condition, for both boys and girls guessing accounted for 35 to 55% of the performance at ages 4 and 5, and less than 10% of the performance at 6 and 8. Both sexes used subclass comparison between 35 and 50% of the time at 4 and 5, and less than 5% of the time at age 8. At age 6, however, girls used subclass comparison about 30% of the time, and boys about 60% of the time.

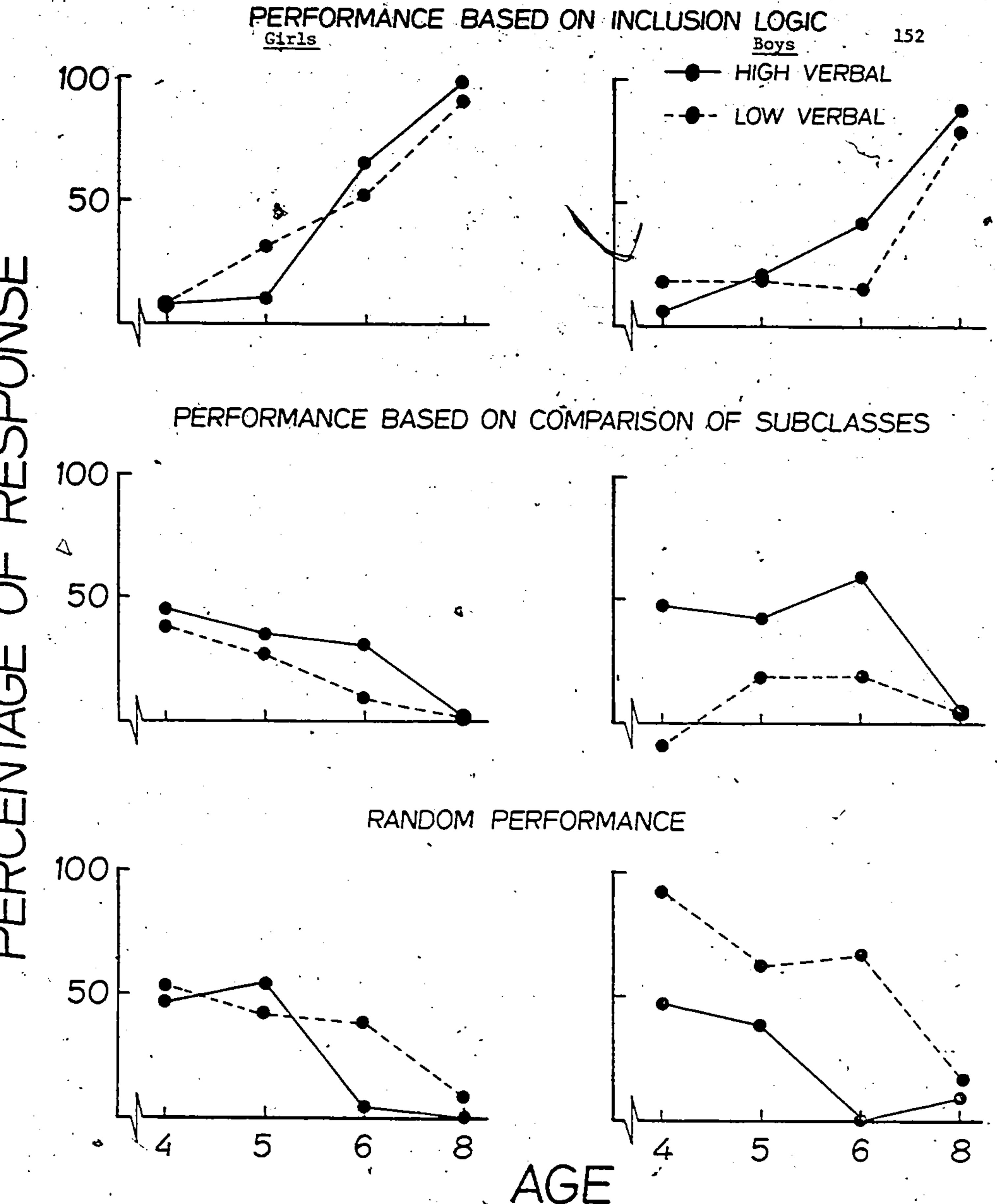
Table 15

Estimated Components of Response to Inclusion Question for Boys and Girls

		Age 4	5	6	8
		GIRLS			
Standard	X	.08	.10	.65	.98
	Y	.45	.35	.31	.02
	Z	.47	.57	.04	.00
Reduced Language	X	.08	.31	.52	.90
	Y	.38	.27	.10	.02
	Z	.53	.42	.38	.08
		BOYS			
Standard	X	.06	.21	.42	.88
	Y	.47	.42	.58	.04
	Z	.47	.38	.00	.08
Reduced Language	X	.18	.19	.15	.79
	Y	-.09	.19	.19	.04
	Z	.92	.63	.67	.17

Note: X: response based on inclusion logic
Y: response based on subclass comparison
Z: response based on guessing

Estimated Components of Inclusion Response for Girls and Boys*



* Each data point represents 8 Ss, 6 responses per S

Figure 12

In the Reduced Language condition, girls used subclass comparison between 25 and 40% of the time at ages 4 and 5, and boys used this strategy less than 20% at the same ages. At 6 and 8, boys used subclass comparison slightly more than girls, but less than 20% of the performance at these ages was based on subclass comparison. There was a larger sex difference in use of guessing in the Reduced Language condition: boys used guessing more than girls at every age level. At 6, boys guessed about two-thirds of the time and girls about three-eights of the time in the Reduced Language condition.

It appears that in this study the sex difference in inclusion performance at age 6 was because the boys used non-inclusion strategies more than girls. In the Standard condition boys were more likely than girls to compare the subclasses. To some extent this was also true in the Reduced Language condition, but a larger effect at this age was that the boys were guessing more than the girls in the Reduced Language condition.

Since this is the only sex difference found in the work reported in this thesis, and since there are no reports of sex differences in the inclusion literature, I suggest it is premature to draw general conclusions. The mathematical model analysis will make it possible in future research to examine strategies used by boys and girls in the inclusion question.

Better Verbal Response in the Reduced Language Condition at Age 4.

Four-year-olds gave significantly more correct verbal response to the Reduced Language inclusion question than to the Standard question. This difference in response decreased at age 5, and at 6 and

8 there was about the same amount of correct response in the two language conditions (see Figure 10). Thus the results obtained in the two conditions show a systematic pattern over age, and the difference at age 4 is consistent with the notion that younger children prefer non-verbal representation such as that used in the Reduced Language condition. We might conclude that young children are better able to show inclusion logic when symbols are used to identify classes. But we are forced to a very different interpretation by the analysis of estimated components of response (Figure 13, Table 16). This analysis indicates that at age 4 there was no difference between the two question forms with regard to amount of performance based on inclusion logic, nor does there appear to have been a difference at ages 5 and 8. If there was a difference in use of inclusion logic it was at 6, when children's responses were more likely to be based on inclusion logic in the Standard condition. That is not to say that there were no differences between the two language conditions with regard to the strategic bases of response. Across the ages studied the children were more likely to use subclass comparison in the Standard condition and to guess in the Reduced Language condition, with both of these non-inclusion strategies approaching zero at age 8.

At 4, when there was a significant difference in correct response, the children were about equally likely to be guessing or comparing the subclasses in the Standard condition. In the Reduced Language condition the four-year-olds were using subclass comparison only about 15% of the time, and they were guessing about three-quarters of the time. These children gave fewer wrong answers to the Reduced

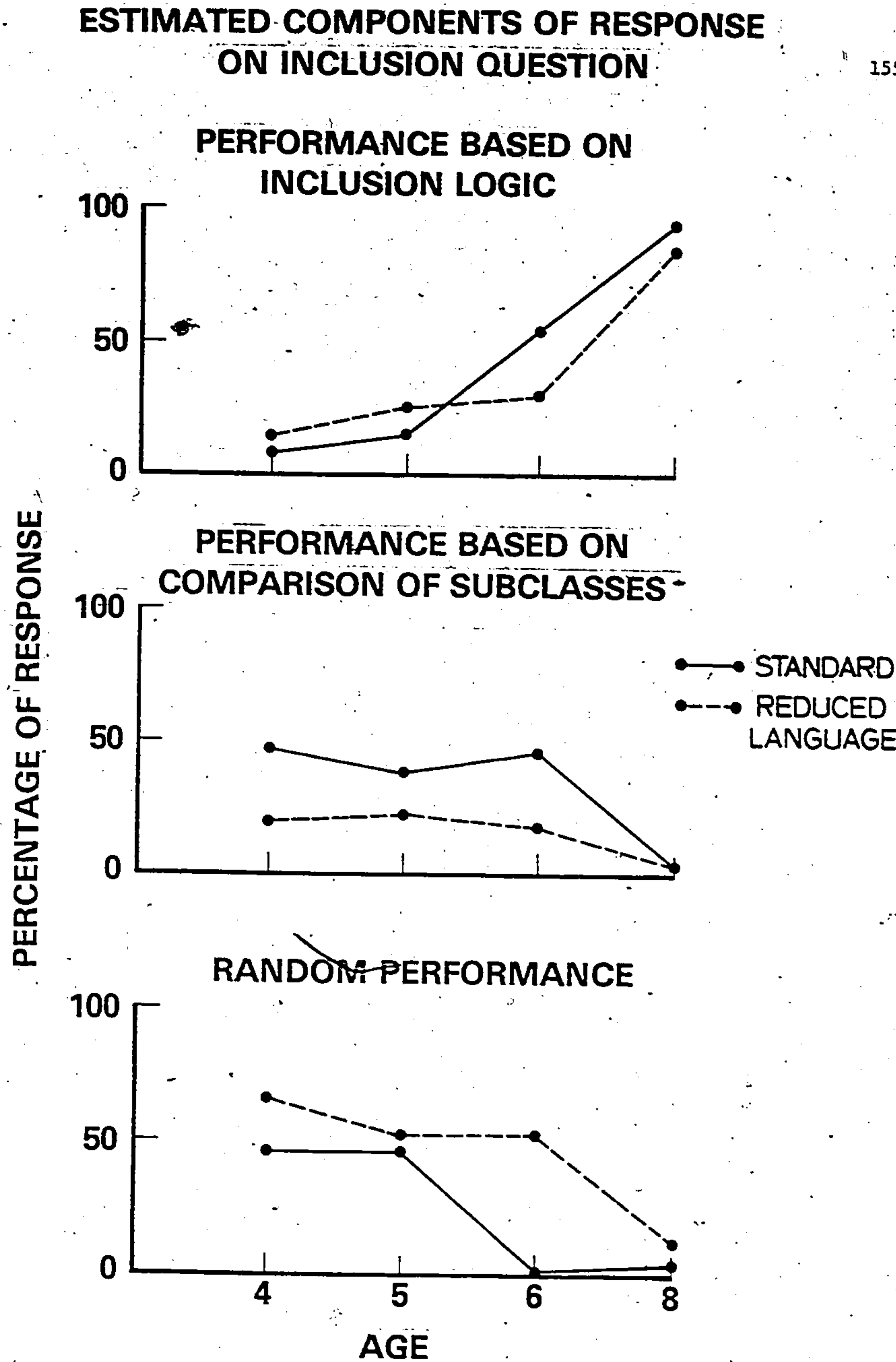

ESTIMATED COMPONENTS OF RESPONSE ON INCLUSION QUESTION
PERFORMANCE BASED ON INCLUSION LOGIC
PERFORMANCE BASED ON COMPARISON OF SUBCLASSES
RANDOM PERFORMANCE
PERCENTAGE OF RESPONSE
100
50
0
STANDARD
REDUCED LANGUAGE
4
5
6
8
AGE

Figure 13

Table 16

Estimated Components of Response to Inclusion Question in the Two Language Conditions

	Age 4	5	6	8
	Standard Condition			
X	.07	.16	.53	.93
Y	.47	.38	.45	.03
Z	.47	.46	.02	.04
	Reduced Language Condition			
X	.12	.25	.33	.85
Y	.15	.23	.15	.03
Z	.73	.52	.52	.12

Note: X: response based on inclusion logic
Y: response based on subclass comparison
Z: response based on guessing

Language question because they were less likely to use subclass comparison, which always produces a wrong answer, and they were more likely to guess which results in wrong answers only half the time.

Summary of Reduced Language Results

To sum up, when language is reduced in the Piagetian inclusion question by using visual symbols to identify the classes children are more likely to use a guessing strategy, and they tend to use this strategy longer than with the standard Piagetian question. In this study virtually no performance in the Standard condition was based on guessing from the age of 6 on, but in the Reduced Language condition the six-year-olds were guessing about half of the time. On the other hand, the children were more likely to use subclass comparison strategy with the standard Piagetian question than with the reduced language form of that question.

This is consistent with the position that the linguistic form of the Piagetian question predisposes to disjunctive comparison, the comparison of subclasses. However this effect of the Piagetian question is not caused by the conjunction "or", which typically joins exclusive classes. If the word "or" were responsible for use of subclass comparison, then there would also be high use of that strategy in the reduced language form "Are there more like this (subclass symbol) or more like this (superordinate class symbol)?" That did not occur in this experiment. Given these results it appears that the naming, or verbal labelling, of classes in the Piagetian question leads to misinterpretation of the question as a request for disjunctive comparison. When visual symbols identify classes this misinterpretation is less likely.

These results do not support the contention made at the beginning of the study that the Piagetian question may be more difficult for children than a question in which language is reduced. At least in this study this was not the case if, by more difficult, we mean that the question delays the demonstration of inclusion logic. When response was partitioned into underlying components, the amount of inclusion-based response was only trivially more for four- and five-year-olds in the Reduced Language question and six-year-olds appeared more likely to use inclusion logic in the Standard question than in the Reduced Language question.

These findings do challenge one aspect of Piagetian theory. The Piagetians correctly observe that young children systematically give wrong answers to the inclusion question because they are comparing the subclasses. Their interpretation of this is that when children are not able to make the simultaneous comparison of superordinate class and included subclass, the children resort to subclass comparison. In other words, Piagetians assert that comparison of subclasses results from logical deficit. The results presented here indicate that this is not the case, but rather that error based on subclass comparison is primarily an effect of the language used in the inclusion question.

Conclusion

By using the mathematical analysis of the inclusion model, we can see that the sex differences in inclusion performance at age 6 are a result of different use of strategy by girls and boys. More important, use of the model prevents erroneous interpretation of the language condition results. Four-year-olds gave more right answers to the reduced language question than to the standard question because of different use of strategy, but the superior performance with the reduced language question was not based on greater use of inclusion logic.

Chapter 9

General Discussion

The object of this research was to examine the effect that the linguistic form of the class inclusion question has on children's response to that question. This was done by comparing different language forms of the inclusion question, and by reducing language in the standard Piagetian question. An attempt was made to determine the strategies that children use at different ages with different forms of the inclusion question.

Different Language Forms: Siegel's and Piaget's

The standard Piagetian question, "Are there more smarties or more candies?", was compared to Siegel's question in the form, "Do you want to have the smarties or the candies?" Siegel, McCabe, Brand and Matthews (1977) reported that three- and four-year-olds gave more correct inclusion answers to Siegel's question than to Piaget's. Considering only verbal response, the results in Experiment 1 replicate that finding for four-year-olds but not for three-year-olds. However, this does not mean that the four-year-olds were demonstrating more inclusion logic with Siegel's question. On the standard form of Siegel's question when children named the superordinate class they often took only the smaller subclass which had not been specified in the question, and this was more likely to happen when the children's less preferred subset was used for the larger subclass in the inclusion question. For example, offered smarties or candies, the children said candies and then put in their cup only the candies they liked best, the jelly beans. In other words, performance on the standard form of Siegel's question was affected by use of non-inclusion strategies, and the higher level of correct response was artifactual, not a matter of inclusion logic. On the other hand, some of the error response on Siegel's question may be artifactual as well. Some of the children might have chosen the larger subclass because those were the items that they wanted, not because they were incapable of making the inclusion comparison.

In contrast to these results with the standard inclusion questions, when the word "all" was used in Experiment 1 questions the children did demonstrate more inclusion logic with Siegel's question than with Piaget's. With the All form, when the three- and four-year-olds gave the correct answer to Siegel's question they usually made the consistent behavioural response, putting the superordinate class in the cup, and the behavioural response was always consistent with correct verbal response after age 4. With the All form of Siegel's question, when the children gave correct verbal answers they meant those answers, and so the superiority of verbal response on Siegel's task in the All condition appears to be based on increased demonstration of inclusion logic. However, the results obtained in Experiment 2 are not consistent with this finding, because in this study children did <u>not</u> do better on Siegel's question, whether "all" was used in the question or not. This difference in results may have been caused by procedural differences in the two studies. Experiment 2 was not designed to compare Piaget's and Siegel's questions, and so each child was asked only two of each of these questions, and the inclusion task was preceded by two other logical tasks. In Experiment 1 there were no other logical tasks, and each child was asked more questions, so these results are more likely to be reliable than those of Experiment 2. Furthermore, the behavioural measure was not taken for the Piagetian questions in Experiment 1 but it was used in Experiment 2, and this change in procedure may have increased the percentage of correct response by making the motivational character of the Piagetian question more like that of Siegel's question. Putting the chosen items in the experimenter's cup may have increased the children's interest in the question. This possibility is supported by the observation that at every age level children gave more correct answers to the standard form of Piaget's question in Experiment 2, in which the behavioural measure was taken, than they did in Experiment 1, which did not include the behavioural measure (see Figure 14).

The discrepancy between Siegel's finding and mine with regard to three-year-olds may be partly accounted for by the fact that many of the three-year-olds in Experiment 1 used a recency strategy on Piaget's question. This strategy raised their performance close to 50%, so that there was less difference between performance on Siegel's and Piaget's questions than would be the case if the children had been using recency strategy less and subclass comparison more on the Piagetian question. In addition there were various differences in the two studies which may have affected the results obtained with all age groups. For example the question form used in my research was, "Do you want to have...", while Siegel et al. used "Do you want to eat...", and the two studies also differed with regard to number of questions asked each child, and variety of stimulus materials.

I suggested that the behavioural measure used in Experiment 1 might have affected inclusion performance on Piaget's question, and it may have done so on Siegel's question as well. In fact, both of the methodological innovations in Experiment 1, the behavioural measure and the control for subset preference, may have changed inclusion response. Asking children which subset they like better might focus their interest on the matter of preference, resulting in greater use of the "other subclass" strategy in which children name the superordinate class in order to get the more preferred smaller subclass. Because children were not asked which subset they preferred in the Siegel et al. study, perhaps they were not using the "other subclass" strategy in that study, but were in fact choosing the superordinate class, all of the items. Similarly, taking the behavioural measure by asking the children to put their chosen items in the cup may affect their responses on all subsequent inclusion questions. They may have perceived that the behavioural instruction was a test of their answer, and that perception may have resulted in lack of confidence in their initial response strategy. If so, some children may have switched from using inclusion logic to choosing the smaller, more preferred subclass.

All of the foregoing says that a) I did not replicate the findings of Siegel et al. except with children older than three in the All condition, and b) there are many reasons why this may be so. My results might have been affected by the children's use of the recency strategy on the Piagetian question, and by various changes in experimental procedure, including asking the children for their subset preference and taking the behavioural measure. It is also possible that Siegel's findings reflect unknown artifactual effects. The most reasonable conclusion, I think, is that the comparison of Siegel's and Piaget's question deserves further study. An appropriate experiment should include three-year-olds, since Siegel et al. report a difference with this age group, and each child should be asked several questions of both types, Piaget and Siegel. Half of the children should receive the behavioural request, and the other half should not, to test possible effects of the behavioural measure on inclusion performance. Similarly, the effect of asking for subset preference should be tested. In addition, for both Piaget's and Siegel's inclusion questions, the children should also be asked the smaller subclass comparison question, so that those responses are available for analysis with the mathematical model.

Different Language Forms: Use of the Word "All"

Another comparison of linguistic forms in this work was between the standard inclusion question and the same question with "all" modifying the superordinate class, in the Piagetian case, "Are there more smarties or more candies?" compared to "Are there more smarties or more of all the candies?" For Siegel's question the two forms were "Do you want to have the smarties or the candies?" and "Do you want to have the smarties or all the candies?" The All questions resulted in more correct inclusion response for ages 3 through 8 and the behavioural data for Siegel's question indicated that children in the All condition understood their correct verbal response better than did the children in the Standard condition.

It seemed possible that at some ages the better performance with "all" was not based on inclusion logic, but rather on a non-inclusion strategy, and Experiment 2 was designed to test this possibility. If children are choosing the class modified by "all", then when the subclass comparison question is "...more A or more of all the A'," they should be more likely to choose A' both verbally and behaviourally. If children are simply choosing all of the items without making any comparison, then on this subclass comparison question they should choose all the items, B, if not verbally then behaviourally; and they should do as well on the Double All inclusion question, "...more of all the A or more of all the B" as on the All question, "...more A or more of all the B." None of these predictions was supported, and so the results indicated that children over the age of 3 were not relying on either of these two non-inclusion strategies. Performance is better when "all" is used in the question because children can understand that question better than the usual linguistic form. This conclusion is consistent with Winer's (1974) observation that the usual form of the inclusion question is often interpreted as asking for disjunctive, not inclusive, comparison. In normal conversation, to refer to both a superordinate class and an included subclass, we typically modify the superordinate with "all". We do this to make clear that we mean the whole superordinate, not the superordinate minus the subordinate, that is "the others".

It appears, then, that the standard Piagetian question invites disjunctive comparison, and so some of the error obtained is based on misunderstanding that question, not on logical inability. This interpretation is supported by the results obtained when language was reduced in the Piagetian question.

Reduction of Language

It was pointed out in Chapter 1 that while there have been many studies which reduced language in the assessment of conservation, seriation, and transitivity only one experiment has reduced language in the assessment of classification, and that experiment was not very successful. There are different positions about the relationship of language and thought. Many current theorists, including the Piagetians, claim that early classification functioning precedes accompanying language, and there is a reasonable amount of evidence to support this position. Another point of view is that there is some degree of structure inherent in the natural world, and to the extent that this is so categorization is defined by that structure, and is independent of language. In either case it seemed worthwhile to try to reduce language in the assessment of classification. My approach to reducing language focussed on the identification of classes in the inclusion question, because that is an important aspect of classification operations. Some sort of imaged or iconic representation is probably preferred by young children, and so I used visual symbols to represent the subclasses and the superordinate class in the inclusion question. The only difference in assessment conditions was that verbal labels, or names, were used in the standard question, and visual symbols were used in the reduced language question.

Four-year-olds gave significantly more correct verbal response in the reduced verbal condition than they did in the standard question condition. This difference in performance decreased with the five-year-olds, and at 6 and 8 correct response was virtually the same in the two language conditions. This finding is consistent with the idea that younger children prefer perceptually based representation, but that language becomes more important as the children get older.

Further Interpretation Based on the Mathematical Model

The mathematical model analysis indicated that the difference in verbal response between the standard form and the reduced language form was not based on increased demonstration of inclusion logic when language was reduced, but rather reflected different use of non-inclusion strategies in the two question conditions. When symbols were used to identify classes a higher proportion of response was based on guessing than in the standard question form, and this difference was particularly large at ages 4 and 6. In the standard question, with words identifying classes, the children were more likely to compare the subclasses, and in comparison to the reduced language question the difference between the two conditions in use of this strategy was greatest at 4 and 6. The percentage of response based on inclusion logic was about the same in the two language conditions, except at age 6 when about 53% of the children in the standard question condition were using inclusion logic, and only about 33% of the children in the reduced language condition did so.

In future research three-year-olds should be included in a reduced language study of class inclusion, to see if their performance and use of strategy is consistent with that of the four-year-olds. It would also be interesting to do a study like this with concrete objects for stimulus materials, in addition to geometric figures. It is more difficult to devise simple visual symbols for real life objects than for abstract geometric items which vary only on the independent dimensions of shape and color, but the results would be more closely related to real life classification functioning. Children deal with candy and drinking straws more than they deal with triangles and abstract "blue things". If the type of stimulus content affected the results of a reduced language study of inclusion, I expect that the difference might be that visual symbol identification of real life classes is more difficult for children than symbolic identification of geometric figure classes. If so, demonstration of inclusion competence would be delayed with real life classes in the reduced language condition, and probably the children would rely longer on a guessing strategy.

Contributions to the Understanding of Language Effects

In the introductory sections I suggested that some linguistic forms of the inclusion question may be hard for children to understand, in ways that have nothing to do with classification ability. I also suggested that varying language in inclusion assessment might provide a better understanding of the development of inclusion ability, and might even add to our understanding of the relationship of language and thought.

Halford (1972) spoke to the issue of question comprehensibility in discussing conservation training studies, including those which used verbal training. He concluded that the evidence was weak that training

> has actually changed the operations of child thought. It is equally likely that the training procedure served to increase the child's knowledge of the conservation situation. The child would then be better able to see that the task was an instance of an operational concept which he had already attained, but the relevance of which he had not seen (p.186).

Halford is saying that changing the assessment of logical capability by using a training intervention may facilitate children's understanding of what is being asked of them. This research has shown that changing inclusion assessment by varying language also affects children's understanding of the inclusion question. This was most clearly demonstrated by use of the word "all" in the question. The children were able to understand this question, and so were more likely to give correct inclusion responses. Increased question comprehensibility may also be involved in the better performance observed on Siegel's question compared to Piaget's, in the All condition of Experiment 1. The contention that the Piagetian question is misleading because it implies disjunctive comparison was supported by the analysis of Experiment 3 results, in which children were more likely to compare subclasses to each other with the standard Piagetian question.

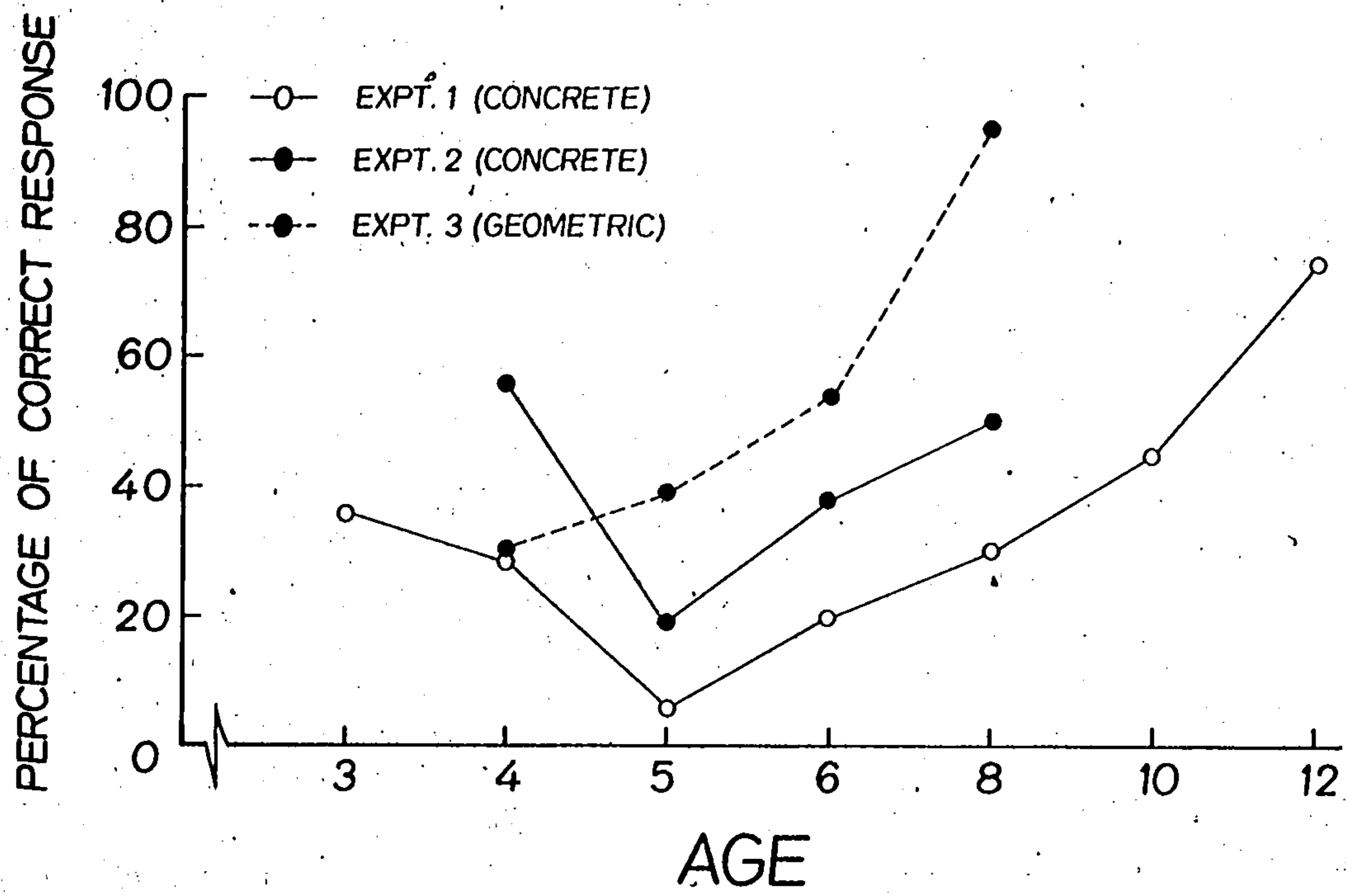
I cannot draw any conclusions from these studies about the structural, or causal, relationship of language and classification operations in development. There is information here about the functional relationship of language and classification, in the sense that we learned about some effects of language on children's use of strategy in class inclusion tasks. The analysis of these data has shown that use of strategy and effects of language on strategy change with age, and so we have a better understanding of the development of inclusion skills. We know that before mastery is attained children are likely to use a subclass comparison strategy with the standard Piagetian question, and they are likely to guess when that question is changed by using visual symbols to identify the classes. At early ages children are more likely to use a strategy based on inclusion logic when the word "all" is used in the inclusion question, and this is particularly true with Siegel's inclusion question.

Effects from type of Stimulus Material

One of the interesting side points of this work is the effect of type of stimulus material on inclusion performance. In Experiments 1 and 2 concrete objects were used, while in Experiment 3 cut-outs of geometric figures were used for the inclusion questions. In each of these three experiments, one of the inclusion forms given the children was the standard Piagetian question, "Are there more A or more B?" The results for this question for Experiments 1, 2, and 3 are shown in Figure 14. At age 4 the percentage of correct responses was highest in Experiment 2, but for children 5 through 8 the highest performance was obtained for the inclusion questions about geometric figures in Experiment 3. The difference between this performance with geometric figures and the performance with concrete objects in the other two cases is quite striking. At age 6, children gave 54% correct responses in Experiment 3, 38% in Experiment 2, and 20% in Experiment 1. At 8, the percentage of correct responses was 95% in Experiment 3, 50% in Experiment 2, and 30% in Experiment 1.

Many factors may cause differences in results between studies,

Inclusion Performance with Concrete Objects and Geometric Figures
In Three Studies*



* Results are shown for the standard Piagetian question in each study.

Figure 14

including differences in samples of children, whether only standard Piagetian questions are asked or whether these questions are alternated with other inclusion forms, whether other experimental tasks precede the inclusion question as in Experiment 2. However, in this research it seems that type of stimulus material caused a very strong effect for children 5 through 8: children at these ages did much better with geometric figures than with real life objects. This finding is consistent with other results in the inclusion literature. Smedslund (1964) found inclusion competence at an earlier age than usual in his study with geometric figures: the seven-year-olds gave 90% correct response. Kohnstamm (1968) used three types of stimulus content in a training study with five-year-olds, and he obtained the best results with drawings of blocks that were essentially geometric figures. Eighty percent of the five-year-olds attained success with these materials. Inclusion competence is demonstrated at a much earlier age when simple, abstract geometric forms are used than when pictures of objects or real objects are used.

<u>Methodological and Analytic Contributions</u>

The contributions of the work reported here with regard to the nature of assessment are both methodological and analytic in nature. For example, the behavioural response measure is useful in evaluating the meaning of children's verbal response. This measure was contributed by a child in the first study who chose what he wanted on Siegel's question and put those items in a paper cup himself, not waiting for the experimenter to do so. If behavioural and verbal response are consistent, we have evidence that children mean what they say, that they understand their answer. When the behavioural and verbal responses are inconsistent, the behavioural response provides an indication of the strategy children are using.

The behavioural response measure in the inclusion question is analogous to information obtained by the Piagetians when they interview children about reasons for answers. The interview approach may be a richer source of information in the sense that children might sometimes provide insights that the experimenter was not looking for. On the other hand, the behavioural response has the advantages of being clearly defined, fast to obtain, and less vulnerable to experimental bias than a discussion between the experimenter and the children.

As mentioned earlier in this chapter, when repeated measures of the inclusion question are taken the behavioural measure has the potential disadvantage of changing the responses that it is meant to check. That is a matter for further empirical investigation. If it does change the nature of responses to the inclusion question, it may do so in a manner which provides further insight into the strategies used in answering inclusion questions.

Another important piece of methodological information concerns use of the children's more preferred or less preferred subset in the inclusion question. Given the two subclasses that make up the superordinate, if the items children like best are used for the larger subclass in the inclusion question they are less likely to give the correct superordinate response. This is not surprising with Siegel's question in which children are asked what they want: they want the items in the larger subclass, and they may consciously decide that they don't want all of the items which includes the smaller subclass. But more or less preferred subset made a difference in response on the Piagetian question as well, and that finding was quite unexpected. There was no reason to think that preference for subsets would affect response to the quantitative comparison asked in the Piagetian question, but it did.

Perhaps when the more preferred subset is used for the larger subclass in the Piagetian inclusion question the children are more likely to pay attention to that class and to use the subclass comparison strategy which leads to wrong responses. This possibility could be tested by using the mathematical model analysis. If this hypothesis is supported, the effect would be similar to that reported in some of the class inclusion studies which examined the salience of the classes in the inclusion question. For example McGarrigle, Grieve, and Hughes (in press) found that there is more correct response when the superordinate class is salient, and less correct response when the subclass is made more salient. In the case of subset preference the salience would not be perceptually or verbally based, but would be a matter of motivational effect.

Because subset preference affects response to the inclusion question, it is important to control for this factor when concrete objects are used as the materials for the inclusion question. There is no evidence in this work about whether subset preference affects performance when abstract geometric figures are used as stimulus items, and one may speculate that blue triangles do not have the same motivational character as smarties. Asking children which items they prefer may affect their response to the inclusion question, and this possibility should be investigated. To study this question, the procedures used in this work could be compared to a condition in which children are asked the series of inclusion questions first, and are asked their preference second.

The final, and perhaps the most important contribution of this research is the analysis of estimated components of inclusion performance provided by the mathematical model. To use this model, the children must be asked to compare the smaller subclass to the superordinate, as well to compare the larger subclass to the superordinate in the inclusion question. The correct and error response to these two questions is then used to estimate the extent to which children are guessing, comparing subclasses, and using inclusion logic.

Just as the behavioural response information was important in understanding the difference in performance between Siegel's and Piaget's questions, the mathematical model analysis was crucial for correct interpretation of performance differences in the reduced language experiment. If only the verbal responses to the inclusion questions were available in that study, we would conclude that more young children show inclusion logic with visual symbols to identify classes, than with verbal labels. We would be wrong. The analysis into underlying components of response shows that performance was higher with the reduced language form because with that question the children were less likely to use the subclass comparison strategy which always leads to error, and they were more likely to guess, which only leads to error half the time. There was no greater use of inclusion logic by the younger children in the reduced language condition.

The mathematical analysis may prove useful in explaining a curious phenomenon. Very often performance is worst for the five-year-olds, so that three- and four-year-olds are at some medium level of correct response, five-year-olds have very low levels of correct response, and then performance improves with increasing age. This pattern of results was not found in Experiment 3, but it was shown in Experiments 1 and 2 (see Figure 14). Five-year-olds are in kindergarten, and when their performance is lower than that of preschool children it is tempting to make undignified comments about the beneficial effects of formal education. It may be, however, that the increased error sometimes shown by five-year-olds does in fact represent developmental progress in the use of strategy. If three- and four-year-olds are likely to guess, and if five-year-olds are guessing less and using subclass comparison more, then the five-year-olds are moving ahead even though subclass comparison leads to a higher percentage of error. Both strategies are "wrong" from an adult point of view, but subclass comparison is a more coherent and systematic approach than random response. The estimated components of response in the standard Piagetian question of Experiment 3 do not support this conjecture, but then the verbal performance function in this case does not show the usual dip at age 5. It is an interesting question that might be examined in future studies.

Final Comments

This work makes it clear that how we ask a question affects the answer we get. That is true not only when we ask a specific question of a child, but it is true of our research endeavors in general. There is a research literature on experimenter effects (e.g., Rosenthal, 1964, 1966; Orne, 1969) and we would do well to review that literature periodically, and remind ourselves of how our view of reality, our predilections, our experience with our own thought processes, affect the questions we choose to study and the way we set about answering those questions.

Some effect of personal bias and academic zeitgeist on research is probably inevitable; our job is to be aware of these effects, and to lessen the constraints of bias on our work. In cognitive developmental research, if we really want to find out about the nature and development of children's systems as Nelson urges, we might pay attention to a wise woman who was working with children early in this century. Maria Montessori (e.g., 1912/1964) said over and over again that we should be still and observe the child.

There are sometimes strong pressures not to observe children as they are, but rather as we are geared to see them. A case in point involves the behavioural measure which was provided by one of the children in this study. I almost missed that contribution. I almost missed it because when the child grabbed the candies and put them in his cup I was busy following my experimental procedure, and his behaviour was not on my program. I was supposed to put those candies in the cup. My initial reaction was along the order of, "What a cheeky kid." But the child's behaviour stayed with me, and somewhat later the realization came: that can tell us something useful. Later still it occurred to me to consider the situation from the child's point of view. (Those familiar with Piagetian theory will recall that egocentrism is characterized by centering on oneself, not considering the perspective of others, and remaining relatively unaffected by information from others.) I was asking the child what he wanted to have, and we were far enough along in the experimental procedure that he knew he did in fact receive his choice. I asked the next question, he answered, and he took his candies - a reasonable, one might even say logical, sequence of events. We cannot know, of course, what other information the children in these studies offered that I did not receive.

Cognitive developmental research will be best served by methodology which includes both naturalistic and experimental approaches. Naturalistic and observational techniques contribute richness of information, and validity in the sense that the data obtained are related to human experience outside the experimental paradigm. Experimental techniques allow a rigor and control that is not possible in field work, and mathematical analysis of the measures gathered experimentally provides a check on the validity of interpretations and conclusions. Our research programs should take advantage of both approaches.

The idea that naturalistic inquiry is important to developmental psychology is not new. The shadow of Jean Piaget is cast upon the work reported here, as it is upon much of the work currently done in cognitive development. Studies which attempt to disprove some aspect of Piagetian theory and studies which seek to confirm Piagetian theory both offer testimony to the impact this man has had on cognitive developmental psychology. And, both kinds of studies may extend the contributions of Piaget, Inhelder, and their co-workers.

Many years ago Piaget played marbles with children in Geneva, and he talked with the children about the rules of the game, where the rules came from, whether they could be changed (Piaget, 1932/1948). He observed children's language with each other before tape recorders were available (Piaget, 1926/1952). This information obtained from observation and the clinical interview method enriched our understanding of development. No doubt there is some Type II error, and possibly other kinds of error as well, in the Piagetian work, and experimental studies can serve to uncover some of that error and to refine the conclusions drawn. The reverse is true as well: observations of children in natural situations can inform and guide experimental research.

REFERENCES

Ahr, P.R. & Youniss, J. Reasons for failure on the class inclusion problem. Child Development, 1970, 41: 131-143.

Anastasi, A. Psychological testing. New York: Macmillan, 1976.

Anglin, J. Word, object, and conceptual development. New York: W.W. Norton, 1977.

Beilin, H. Learning and operational convergence in logical thought development. Journal of Experimental Child Psychology, 1965, 2: 319-339.

Beilin, H. & Kagan, J. Pluralization rules and the conceptualization of number. Developmental Psychology, 1969, 1: 692-706.

Best, B.J. Classificatory development in deaf children: Research on language and cognitive development. Occasional Paper #15, Department of Health, Education, and Welfare, April, 1973.

Blank, M. Cognitive functions of language in the preschool years. Developmental Psychology, 1974, 10(2): 229-245.

Blank, M. & Altman, L.D. Effects of stimulus modality and task complexity on discrimination and reversal learning in preschool children. Journal of Experimental Child Psychology, 1968, 6: 598-606.

Blank, M. & Bridger, W.H. Cross-modal transfer in nursery school children. Journal of Comparative and Physiological Psychology, 1964, 58: 277-282.

Braine, M.D.S. Piaget on reasoning: A methodological critique and alternative proposals. Monographs of the Society for Research in Child Development, 1962, 27(2): 41-61.

Brainerd, C.J. Judgments and explanations as criteria for the presence of cognitive structures. Psychological Bulletin, 1973, 79(3): 172-179.

Brainerd, C.J. Training and transfer of transitivity, conservation, and class inclusion of length. Child Development, 1974, 45: 324-334.

Brainerd, C.J. & Kaszor, P. An analysis of two proposed sources of children's class inclusion errors. Developmental Psychology, 1974 10(5): 633-643.

Brooks, L. Non-analytic concept formation and memory for instances. Working paper for S.S.R.C. conference on human categorization, 1976.

Brown, R. Language and categories. In Bruner, Goodnow, Austin (Eds.), A study of thinking. New York: Wiley, 1956.

Brown, R. Psycholinguistics. Toronto: The Free Press, 1972.

Brown, R. & Lenneberg, E.H. A study in language and cognition. Journal of Abnormal and Social Psychology, 1954, 49: 454-462.

Bruner, J.S. On cognitive growth. In J.S. Bruner, R.R. Olver, & P.M. Greenfield (Eds.), Studies in cognitive growth. New York: Wiley, 1966.

Bruner, J.S. The course of cognitive growth. American Psychologist, 1964, 19: 1-15.

Bruner, J.S., Goodnow, J.J. and Austin, G.A. A study of thinking. New York: John Wiley, 1956.

Carroll, J.B. & Casagrande, J.B. The function of language classifications in behavior. In E.E. Maccoby, T.M. Newcomb, & E.L. Hartley (Eds.), Readings in social psychology. (3rd ed.) New York: Holt, Rinehart and Winston, 1958.

Carson, M.T. & Abrahamson, A. Some members are more equal than others: The effect of semantic typicality on class-inclusion performance. Child Development, 1976, 47: 1186-1190.

Clark, E.V. Strategies and the mapping problem in first language acquisition. In J. Macnamara (Ed.), Language learning and thought. New York: Academic Press, 1977.

Cole, M. & Scribner, S. Culture and thought. New York, London: John Wiley & Sons, 1974.

Dale, P.S. Language development. Hinsdale, Ill.: The Dryden Press, 1972.

Exhibit A: IQ trial. APA Monitor, 1977, 8(12): 4-5.

Furth, H.G. Linguistic deficiency and thinking: Research with deaf subjects 1964-1969. Psychological Bulletin, 1971, 76(1): 58-72.

Furth, H.G. & Milgram, N.A. The influence of language on classification: A theoretical model applied to normal, retarded, and deaf children. Genetic Psychology Monographs, 1965, 72: 317-351.

Halford, G.S. The impact of Piaget on psychology in the seventies. In P.C. Dodwell (Ed.), New horizons in psychology. Harmondsworth, England: Penguin, 1972.

Heider, E.R. "Focal" color areas and the development of color names. Developmental Psychology, 1971, 4: 447-455.

Hollos, M. & Cowan, P.A. Social isolation and cognitive development: Logical operations and role-taking abilities in three Norwegian social settings. Child Development, 1973, 44: 630-641.

Inhelder, B. & Piaget, J. [The early growth of logic in the child.] (Translated by E.A. Lunzer & D. Papert). New York: W.W. Norton, 1969. (Originally published, 1964).

Isen, A.M., Riley, C.A., Tucker, T., & Trabasso, T. The facilitation of class inclusion by use of multiple comparisons and two-class perceptual displays. Paper presented at the Society for Research in Child Development, Denver, Colorado, April, 1975.

Kalil, K., Youssef, Z. & Lerner, R.M. Class-inclusion failure: Cognitive deficit or misleading reference? Child Development, 1974, 45: 1122-1125.

Klahr, D. & Wallace, J.G. Cognitive development: An information-processing view. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1976.

Koff, E. & Luria, Z. Concept and language: The comparative relation. Paper presented at Society for Research in Child Development meeting, Philadelphia, 1973.

Kofsky, E. A scalogram study of classificatory development. In I.E. Sigel and F.M. Hooper, (Eds.), Logical thinking in children: Research based on Piaget's theory. New York, Toronto, London: Holt, Rinehart & Winston, 1968.

Kohnstamm, G.A. An evaluation of part of Piaget's theory. In I.E. Sigel & F.H. Hooper (Eds.), Logical thinking in children: Research based on Biaget's theory. New York: Holt, Rinehart & Winston, 1968.

Lantz, D. & Lenneberg, E.H. Verbal communication and color memory in the deaf and hearing. Child Development, 1966, 37: 765-799.

Lantz, D. & Stefflre, V. Language and cognition revisited. Journal of Abnormal and Social Psychology, 1964, 69: 472-481.

Lawson, G.A. Number, language, and cognitive development. Unpublished doctoral dissertation, McMaster University, 1978.

Lenneberg, E.H. Color naming, color recognition, color discrimination: A reappraisal. Perceptual and Motor Skills, 1961, 12: 375-382.

Luria, A.R. & Yodovich, F.I. [Speech and the development of mental processes in the child.] Translated by O. Kovasc and J. Simon. London: Staples Press, 1968. (Originally published, 1959.)

Macnamara, J. On the relation between language learning and thought. In J. Macnamara (Ed.), Language learning and thought. New York: Academic Press, 1977.

Markman, E.M. The facilitation of part-whole comparisons by use of the collective noun "family". Child Development, 1973, 44: 837-840.

Markman, E.M. Empirical versus logical solutions to part-whole comparison problems concerning classes and collections. Child Development, 1978, 49: 168-177.

Markman, E.M. & Seibert, J. Classes and collections: Internal organization and resulting holistic properties. Cognitive Psychology, 1976, 8: 561-577.

McGarrigle, J., Grieve, R., & Hughes, M. Interpreting inclusion: A contribution to the study of the child's cognitive and linguistic development. Child Development, in press.

Miller, S.A. Nonverbal assessment of Piagetian concepts. Psychological Bulletin, 1976, 83: 405-430.

Miller, S.A., Shelton, J., & Flavell, J.H. A test of Luria's hypothesis concerning the development of verbal self-regulation. Child Development, 1970, 41: 651-665.

Montessori, M. The Montessori method. New York: Schocken, 1964. (Originally published, 1912).

Moore, T.E. & Harris, A.E. Language and thought in Piagetian theory. In L.S. Siegel and C.J. Brainerd (Eds.), Alternatives to Piaget. New York: Academic Press, 1978.

Nelson, K. The conceptual basis for naming. In J. Macnamara (Ed.), Language learning and thought. New York: Academic Press, 1977.

Oléron, P. Recherches sur le developpement mental des sourds-muets. Contribution a l'etude du probleme. Paris: Centre National de la Recherche Scientifique, 1957. Cited in M. Blank, Cognitive functions of language in the preschool years. Developmental Psychology, 1974, 10(2): 229-245.

Olson, D.R. The contexts of language acquisition. In J. Macnamara (Ed.), Language learning and thought. New York: Academic Press, 1977.

Orne, M.T. Demand characteristics and the concept of quasi-controls. In R. Rosenthal and R.L. Rosnow (Eds.), Artifact in behavioral research. New York: Academic Press, 1969.

Penn, J.M. Linguistic relativity versus innate ideas. Paris and the Hague: Mouton, 1972.

Piaget, J. The language and thought of the child. London: Routledge and Kegan Paul, 1952. (Originally published, 1926).

Piaget, J. The moral judgment of the child. Glencoe, Ill.: Free Press, 1948. (Originally published, 1932).

Piaget, J. [Play, dreams and imitation in childhood.] (C. Gattegno and F.M. Hodgson, trans.) New York: W.W. Norton, 1962.

Piaget, J. [The child and reality] Translated by A. Rosin. New York: Penguin Books, 1976. (Originally published, 1972).

Rosch, E. On the internal structure of perceptual and semantic categories. In T.E. Moore (Ed.), Cognitive development and the acquisition of language. New York and London: Academic Press, 1973.

Rosch, E. & Mervis, C.B. Family resemblances: Studies in the internal structure of categories. Cognitive Psychology, 1975, 7: 573-605.

Rosch, E., Mervis, C.B., Gray, W.D., Johnson, D.M., & Boyes-Braem, P. Basic objects in natural categories. Cognitive Psychology, 1976, 8: 382-439.

Rosch, E. Human categorization. In N. Warren (Ed.), Advances in cross-cultural psychology (Vol. 1). London: Academic Press, 1977.

Rosenthal, R. The effect of the experimenter on the results of psychological research. In B.A. Maher (Ed.), Progress in experimental personality research. Vol. I. New York: Academic Press, 1964.

Rosenthal, R. Experimenter effects in behavioral research. New York: Appleton-Century-Crofts, 1966.

Saravo, A. & Gollin, E.G. Oddity learning and learning sets in children. Journal of Experimental Child Psychology, 1969, 7: 541-552.

Scholnick, E. & Adams, M. Relationships between language and cognitive skills: Passive voice comprehension, backward repetition, and matrix permutation. Child Development, 1973, 44: 741-746.

Sheppard, J.L. Conservation of part and whole in the acquisition of class inclusion. Child Development, 1973, 44: 380-383.

Siegel, L.S. The relationship of language and thought in the young child. Unpublished manuscript, McMaster University, 1976.

Siegel, L.S. The relationship of language and thought in the preoperational child. In L.S. Siegel & C.J. Brainerd (Eds.), Alternatives to Piaget. New York: Academic Press, 1978.

Siegel, L.S. & Goldstein, A.G Conservation of number in young children: Recency versus relational response strategies. Developmental Psychology, 1969, 1(2): 128-130.

Siegel, L.S., McCabe, A.C., Brand, J., & Matthews, J. Evidence for the understanding of class inclusion in the preschool child: Linguistic factors and training effects. Unpublished manuscript, McMaster University, 1977.

Sigel, I.E., Roeper, A., & Hooper, F.H. A training procedure for acquisition of Piaget's conservation of quantity: a pilot study and its replication. British Journal of Educational Psychology, 1966, 36: 301-311.

Sinclair-de-Zwart, H. Developmental psycholinguistics. In Elkind, D. & Flavell, J.H., Eds., Studies in cognitive development. New York, London, Toronto: Oxford University Press, 1969.

Sinclair-de-Zwart, H. Language acquisition and cognitive development. In T.E. Moore (Ed.), Cognitive development and the acquisition of language. New York, London: Academic Press, 1973.

Smedslund, J. Concrete reasoning: A study of intellectual development. Monographs of the Society for Research in Child Development, 1964, 29, (2, Serial No. 93).

Springer, S. The development of conservation of number in deaf and hearing children. Paper presented at the annual meeting of the Canadian psychological Association, June, 1976.

Tatarsky, J.H. The influence of dimensional manipulations on class-inclusion performance. Child Development, 1974, 45: 1173-1175.

Trabasso, T., Isen, A.M., Dolecki, P., McLanahan, A.G., Riley, C.A., & Tucker, T. How do children solve class-inclusion problems? To appear in R.S. Siegler (Ed.), Children's thinking: What develops? Hillsdale, N.J.: Lawrence Erlbaum Associates, in press.

Vygotsky, L.S. [Thought and language.] Translated by E. Hanfmann and G. Vakar. Cambridge, Mass: M.I.T. Press, 1962. (Originally published, 1934).

Weil, J. The relationship between time conceptualization and time language in young children. Unpublished doctoral dissertation, The Graduate Center, City University of New York, 1970. Cited by L.S. Siegel, The relationship of language and thought in the preoperational child. In L.S. Siegel & C.J. Brainerd (Eds.), Alternatives to Piaget. New York: Academic Press, 1978.

Whorf, B.L. In J.B. Carroll (Ed.), Language, thought, and reality: Selected writings of Benjamin Lee Whorf. New York: John Wiley & Sons, 1956.

Whorf, B.L. Linguistics as an exact science. In L.F. Dean & K.G. Wilson (Eds.), Essays on language and usage. New York: Oxford University Press, 1959.

Wilkinson, A. Counting strategies and semantic analysis as applied to class inclusion. Cognitive Psychology, 1976, 8: 64-85.

Winer, G.A. An analysis of verbal facilitation of class-inclusion reasoning. Child Development, 1974, 45: 224-227.

Wohlwill, J.F. Responses to class-inclusion questions for verbally and pictorially presented items. Child Development, 1968, 39: 449-465.