DOCUMENT RESUME

ED 097 102 PS 007 344

AUTHOR Shipley, Elizabeth F.
TITLE The Piagetian Class-Inclusion Task: An Alternative Explanation. Technical Report No. 19.
SPONS AGENCY National Inst. of Mental Health (DHEW), Bethesda, Md.
PUB DATE Feb 74
NOTE 68p.

EDRS PRICE MF-$0.75 HC-$3.15 PLUS POSTAGE
DESCRIPTORS Children; *Classification; Cognitive Ability; Cognitive Development; *Conservation (Concept); Cues; *Developmental Tasks; Generalization; Logic; Psycholinguistics; *Research; *Structural Linguistics; Tables (Data)
IDENTIFIERS Class Inclusion Task; Grammatical Constraint Hypothesis; *Piaget (Jean)

ABSTRACT

This study investigated the linguistic components of Piaget's class-inclusion task. First, hierarchical classification is examined from both Piagetian and linguistic theory points of view. Then, two general characteristics of child thinking that relate to the different interpretations of the responses to classification questions are discussed: (1) the tendency to overgeneralize rules, and (2) the lack of analytic aptitude. The considerations of these characteristics lead to the conclusion that children "understand wrongly" in the class-inclusion tasks because of grammatical constraints. In the third section, five research experiments designed to evaluate the grammatical constraint hypothesis are presented: (1) a methodological study to establish two slightly different forms of the class-inclusion question; (2) a study of grammatical aids to correct comparison; (3) a direct test of the child's use of grammatical clues; (4) a study to determine the independence of class-inclusion performance and the ability to compare part and whole in other ways; and (5) a study to examine performance for two types of numerical comparisons. The results of the experiments and some possible conclusions are also discussed. (Author/SDH)

Grant #MH 23505 - The Acquisition of Linguistic Structure

Technical Report XIX

The Piagetian Class-Inclusion Task: An Alternative Explanation

Elizabeth F. Shipley

University of Pennsylvania

February, 1974

Piaget and his coworkers (Inhelder and Piaget, 1964; Piaget, 1965) have studied children's knowledge of hierarchical organization; the general claim of this group is that the ability to think in terms of superordinate and subordinate classes, and to recognize class-inclusion relations, is a function of emergent cognitive organization. Language provides the child with tools for his cognitive activities, in this view, but language knowledge *per se* plays only a minor role in the mastery of "structures" such as hierarchical classification; it then follows that language provides the experimenter with a convenient and legitimate means to probe the child's cognitive competence, just because this cognitive competence is (a) expressible in language and at the same time (b) independent of characteristics of language development. Thus, apparently, the Piagetian group entertains few reservations concerning the child's ability to comprehend those questions used to probe cognitive competence. Any difficulty in answering such questions is attributed to extralinguistic sources.

However, developmental psycholinguists (e.g., Chomsky, 1969; Lasser, 1970) have shown that the child has not yet achieved complete syntactic mastery (i.e., comprehension) of his first language at the age of eight or even ten years. Further, young children are unable to take a "non-empirical" approach to linguistic problems (Osherson and Markman, 1973), to treat language as an object and to attain meta-linguistic awareness of language (Gleitman, Gleitman, and Shipley, 1973;

Cazden, 1972). Whatever the child's comprehension, then, he may have residual difficulties in employing language specifically as a tool for analytic reasoning. Such findings suggest that we should not take for granted that the child will interpret verbal tasks in ways no different from adults; verbal difficulties may masquerade as reasoning deficits. In short, linguistic aspects of cognitive mastery deserve independent examination; we consider here linguistic components of the well-known "class-inclusion task." We shall argue on the basis of both a theoretical analysis of this task and experimentally derived data that linguistic and cognitive structures interact in subtle ways in children's responses to Piaget's verbal probes.

Below, we first (Section A) describe hierarchical classification from the point of veiw both of Piagetian theory and linguistic theory and formulate an account of children's class-inclusion performance. We then (Section B) describe some general characteristics of child thinking that bear on the interpretation of their responses to classification questions, namely a) the tendency to overgeneralize rules and b) the lack of analytic aptitude. Such consideration lead us to a hypothesis that children "understand wrongly" in the class-inclusion tasks because of grammatical constraints that are partially, at least, independent of cognitive deficits. Third (Section C) we provide an experimental evaluation of the grammatical constraint hypothesis; the results are discussed in Section D.

A. Hierarchical Classification

Mastery of hierarchical classification consists of the apprehension of certain relations which, for reference in further discussion, we here define using standard class nomenclature. The discussion of the algebra of classes is based upon a modern algebra text (Birkhoff and Maclane 1958) and the minimal characterization of hierarchical classification is abstracted from a variety of sources in psychology.

By a class is meant a set or group of elements or things. The basic class relation is that of inclusion of one class within another; class A is included in class B when every element of class A is also an element of class B. Three operations may be performed upon classes: union, intersection and complementation. The union of two classes X and Y is the class whose elements are in either X or Y or both. The intersect of two classes X and Y is the class of all elements in both X and Y. The complement of class X is the class of all elements not in class X. The operations of union and intersection are analogous to the arithmetical operations of addition and multiplication.

Classification is the process of assigning elements to classes; a classification system is the result of the process of classification. The defining property of hierarchical classification is that classification occurs at more than one level and that a class of a lower level or rank is included in a class of each higher level. Usually, but not always, two or more classes of the same level, A and A'; are included in the same class B on the level immediately above.

When class A is included in class B in a hierarchical classification system class B is called the <u>superordinate class</u> and class A is called the <u>subordinate class</u> or the <u>subclass</u>. Subclasses at the <u>same level</u> do not overlap - they do not have elements in common; if some element of class B is in subclass A, every element of class B is in either subclass A or in some other subclass of the same level as subclass A. Equivalently, in terms of the operations of the algebra of classes, the union of all subclasses of the same level of class B is class B itself, and the intersect of any two subclasses of the same level contains no elements. Another way of describing the set of subclasses of the same level of class B is to say that they form a <u>partition</u> of class B. Classes at the same level we will call <u>coordinate classes</u>. Thus we could have a collection of coins as a superordinate class and each subclass would contain coins from a different country. Two classes of coins from two different countries would be coordinate classes.

So far we have said nothing about the basis for assigning class membership. Class membership can be specified simply by pointing at an object and specifing that the object is in class A and class B,etc. Such a way of classifying is denotative or extensive. Alternatively, classification can be based upon properties of elements, and class membership is specified by listing properties rather than by pointing to elements. This is connative or intensive classification.

The distinction between connotation and denotation is an important one in philosophy and in psychology. Inhelder and Piaget (1964) and Vygotsky (1962) have provacative discussions of the importance of the distinction in the child's mastery of concepts and classification. However much of their writings appear speculative and further clarification is needed of such notions as "coordination of extension and intension" (Inhelder and Piaget, 1964).

When we examine existing systems of hierarchical classification in natural language or the sciences we see that all have a connative basis for classification, properties are listed which determine class membership. Further, all subclasses at the same level are defined by the same <u>kinds</u> of properties. Thus class B might be a collection of coins which is partitioned into subclasses on the basis of face value, so that class A contains pennies, class A' contains nickles, etc. We do not find hierarchical classification systems analogous to a collection of coins in which one subset consists of pennies and the other subsets at the same level consist of non-pennies grouped by date of minting.

In the discussion below we first (Section 1) examine Piaget's analysis of the emergence of knowledge in this domain; and second (Section 2) we describe the expression of hierarchical classification and class comparison in English; we will hypothesize (Section 3) that certain difficulties children experience in the Piagetian class-inclusion tasks are attributable to linguistic constraints on the comparison of superordinate and subordinate classes.

1. The Piagetian Analysis

Piaget and his co-workers wrote extensively on hierarchical classification in The Child's Conception of Number (Piaget, 1965),originally published in 1941 by Piaget and Szeminska, and later in The Early Growth of Logic in the Child (Inhelder and Piaget, 1964),originally published in 1959. The treatment in the two books is very similar except for the child's performance at an intermediate stage (Stage II) on the so-called class-inclusion problem. We will briefly consider some material common to the two books and then consider the difference.

Piaget claims to take from logic the nature of the adult's organization of concepts. The structure is that of hierarchical classification which is characterized by "norms of reasoning to which the subject himself conforms" (Inhelder and Piaget 1964, p. 48). One of the norms, and the last to be mastered, is class inclusion: "A class A (or A') is included in every higher ranking class which contains all its elements, starting with the closest, B: A=B-A'(or A'=B-A) and AxB=A, which amounts to saying that all A are 'some' B." (1964, p. 48).

Psychologically, for Piaget, class inclusion involves conservation of the whole: ..."in the case of true inclusion, B, the larger class, does not exist only when its constituent parts, A and A', are actually united ... (but)... it continues to encompass them, and it conserves its identity, even when these are dissociated ... be it in space or even in thought ..." (1964, p. 49, 50). Logically, mastery of class-inclusion is dependent upon mastery of reversible operations, namely the addition and subtraction of classes: B=A+A' and A=B-A'. Addition and subtraction of classes correspond to the class operations of union and complementation under the conditions Inhelder and Piaget are discussing.

Mastery of class-inclusion is manifest in two ways: "The conservation of the whole and the quantitative comparison of whole and part are the two essential characteristics of genuine class-inclusion..." (1964, p. 117). These two characteristics have been explored with two behavioral indices. The first index involves the ability to answer correctly and to justify the answer to questions of the form "Are all females adults?"; this is said to indicate mastery of the "all-some" relation. The second index, usually called the class-inclusion task, involves the ability to compare class with subclass, i.e. to answer correctly questions of the form "Are there more children or girls?"

The typical Piagetian class-inclusion task involves presenting the child with five toy dogs and three toy cats and asking "Are there more dogs or animals?" Young children consistently respond "dogs". According to Piaget the young child is unable to "conserve" the class of animals, so he "reduces" the class of animals to cats and replies as if he were asked "Are there more dogs or cats?"

The earlier work, The Child's Conception of Number, appears to place greater emph sis on the logical aspect of mastery, the additive composition of classes, while the later book, The Early Growth of Logic in the Child, appears to place greater emphasis on the psychological aspect of master, the conservation of the whole. (The emphasis on conservation of the whole makes this approach consistent with other Piagetian explanations, e.g. conservation of number, of mass, of volume, etc.). This difference in emphasis probably accounts for the one substantive difference in the the treatment of the class-inclusion task in the two books, namely the performance of the Stage II child. In the earlier book, the Stage I child is said to be able to think simultaneously of part and whole and hence to be unable to respond correctly; the Stage II child can think simultaneously of B and A under special circumstances and hence he can respond correctly but he does so "intuitively", empirically, by counting. Only the Stage III child "grasps immediately that class B is larger than class A because he approaches the problem from the point of view of additive composition." (1965, p. 164). In the later book, with greater emphasis on conservation of the whole, the distinction between an empirical solution and a deductive solution is dropped; only the Stage III child can answer correctly. Evidently if the whole is not conserved then the child cannot be expected to count all elements of the whole.

The distinction between an empirical solution and a deductive solution should be kept in mind because some children do obtain the correct answer in the class-inclusion task by counting although most appear to arrive at the correct answer deductively.

Is Piaget correct in his analysis of his subjects' behavior? A number of questions of interpretation come to mind. For example, how do we know that the child is comparing coordinate parts when he replies "dogs?" The fact that the child replies "dogs" does not prove that he was comparing dogs with cats. What little evidence exists is found in the children's protocols: children occasionally refer to the two subclasses in some way, e.g. by giving the number of individuals in each of the two subclasses or by mentioning the smaller subclass by name, and some children reply "same" when the two subclasses are equal in number (Inhelder and Piaget 1964, Ahr and Youniss, 1970).

If we grant that the child is comparing part and part rather than part and whole as requested in the class-inclusion task, two separate theoretical questions arise: (1) why doesn't the child compare the whole with the part; and (2) why does he compare the part with another part? Any theoretical attempt to explain failure on the class-inclusion task should also explain the children's consistent behavior when they fail. There is no necessary reason why the inability to engage in the process of comparing part and whole should lead to a comparison of component parts. The child could compare the dogs before him with all other animals in the world, or with other dogs, or with the experimenter's fingers, or with any other class of things. If these possibilities seem farfetched, this itself indicates that the reader acknowledges a certain cohesiveness about the group of objects presented: it is "a whole" which does not include such irrelevant things as other animals or experimenter's fingers. Evidently this sense of the group is grasped by children also: the whole may be "destroyed" but the coherence of the parts is sufficient to limit the comparison to those parts. At some level, the child recognizes a hierarchy which consists of a superordinate class (the whole) and subordinate classes (parts of the whole). The particular error of the "non-conserver"[1] in the class-inclusion task provides internal evidence for the existence of this hierarchy.

Piaget accounts for failure in the class-inclusion task by postulating an inability to conserve the whole; he posits a "reduction" of whole to part to explain the consistent error. But does the reduction of the whole B to the part A' follow from Piaget's model? In fact, nothing in the possible properties of non-graphic collections (1964, p. 48) said to be formed by "pre-conservers" implies the reduction of a destroyed B to A'. Further, Inhelder and Piaget (1964, p. 106), mention other possible errors, suggesting that they themselves do not see the reduction error as a necessary consequent of failure to conserve. We must tentatively regard the reduction hypothesis as an *ad hoc* explanation of children's performance.

Some other attempts to explain performance on the class-inclusion task resort to mininterpretation of the instructions as a request for a comparison of subclasses (Wohlwill 1968, Ahr and Youniss 1970, Hayes 1972). However, we still do not know why the misinterpretation occurrs. Hayes (1972) suggests that a request to compare quantity may lead to a set for an empirical solution; but this suggestion does not explain why subclasses are counted and compared rather than the superordinate class and a subclass as requested. Klahr and Wallace (1972) present two information processing models to explain failure which assume that young children cannot count the same element twice; these models do explain why the children make the specific error of comparing subclasses; however, as they deal only with performance based upon counting, these models cannot account for the children who solve the problem deductively.

In the next section we will attempt to explain the child's error in the class-inclusion task by reference to certain syntactic and semantic features of class comparisons.

2. Some linguistic constraints that operate in the class-inclusion task.

In a theoretical discussion of hierarchical organization in the lexicon, Bever and Rosenham (1971) hypothesize certain restrictions on comparative constructions. They state that constructions are grammatical[2] "just in case a comparing noun neither dominates nor is dominated by a compared noun in the Be hierarchy (p. 593)." Thus, the sentences

(1) A gun is more deadly than a pistol.

(2) A pistol is more deadly than a gun.

are ungrammatical because gun dominates pistol in the hierarchy specified by the verb Be (a pistol is a gun), while both

(3) A cannon is more deadly than a pistol.

and

(4) A pistol is more deadly than a cannon.

are grammatical because neither noun dominates the other in the Be hierarchy (neither a cannon is a pistol nor a pistol is a cannon). Notice that the falsity of (4) does not alter its grammatical status.

Let us translate this constraint into the class nomenclature, introduced earlier using these same examples. Gun names the individual elements of a class; pistol names the individual elements of one subclass of gun, cannon names elements of another subclass of gun. The two subclasses, pistol and cannon do not over-lap; they are coordinate classes. Elements identified only as members of the class gun cannot be compared with elements of a subclass of gun; they cannot be compared with pistol or cannon. This restriction on the comparison of elements of a class and one of its subclasses renders ungrammatical a variety of putative comparative constructions, e.g.

(5) Guns are more deadly than pistols.

(6) Which are more deadly: guns or pistols?

The ungrammatical constructions refer to distributive comparisons of classes, i.e. comparisons of properties possessed by each individual element of a class. Relevant to our analysis of the class-inclusion task we conclude: distributive comparisons of class and subclass are ungrammatical.[3]

The examples of grammatical comparative constructions presented by Bever and Rosenham all contain nouns which refer to elements of coordinate classes. Are all grammatical distributive comparisons between coordinate classes? The answer is not obvious.

The fact that distributive comparisons are usually of coordinate classes can be demonstrated by considering open-ended comparisons.

(7) Apples are more tart than __________.

(8) Rats are more frightening than __________.

(9) Airplanes are faster than __________.

One tends to complete such statements with coordinate classes from well-established lexical hierarchies - pears, mice, boats. Less common completions also imply some hierarchy e.g. rats are more frightening than ice-storms among the natural hazards of some challenging environment.

Bur what of a distributive comparison in which the classes are not coordinate in an obvious classification system. For instance:

(10) Cats are smarter than poodles.

Although this statement is completely comprehensible as it stands, it does seem to imply one of two contexts. Either poodles and cats do have coordinate status in some more unusual hierarchy than biological classification, e.g. rat catchers, apartment dwelling pets, etc. or the comparison of interest is between dogs and cats, coordinate classes, and poodles have some special status as a subclass of dogs. In the later case the stated comparison of cats and poodles is elliptical for some more complicated comparison e.g.

(11) Cats are even smarter than poodles, the smartest dogs of all, so cats are smarter than all dogs.

(12) Cats are at least smarter than poodles, the dumbest dogs of all, but that's all one can say for cat's intelligence with respect to dog's intelligence.

(13) Cats are smarter than poodles but we don't know about other dogs.

Of course if we compare any two classes of things, say cats and parades (Are cats more fun than parades?), we are comparing coordinate classes in a trivial sense; cats and parades are two subclasses of the same superordinate class, things we compare. Beyond this trivial comparison, I submit that comparisons such as (10) presuppose either some other pre-existing hierarchy in which the compared classes are coordinate or an implied comparison of coordinate classes. Whether the nature of the process of distributive comparison is such that coordinate classes are always involved in some way, or whether distributive comparisons are frequently of coordinate classes so that we have a strong expectation that coordinate classes will be involved does not matter for our present purposes.

We conclude: distributive comparisons usually involve coordinate classes.

Let us now turn to comparison on the basis of collective properties - properties possessed by a class as a whole. Properties wuch as numerosity, physical extension, weight, monetary value, usefulness, etc. can be the basis of collective comparison. With the exception of numerosity, both collective and distributive comparisons can be made on the basis of these properties. Yet we do not find constructions that are taken to be ambiguous with respect to the kind of comparison, collective or distributive. For example the following are taken distributively, even though, logically, either are possible.

(14) Adults weigh more than children.

(15) Lakes are more extensive than rivers.

(16) Which are more valuable: gold ingots or cut diamonds?

To be interpreted as a collective comparison, a more elaborate form is necessary.

(17) The total weight of all the adults in the word is greater than that of all the children.

(18) All the lakes combined are more extensive than all the rivers.

(19) Which are more valuable: all the gold ingots or all the cut diamonds?

This additional information, this special wording is necessary to signify that the comparison is between classes, e.g. the class all adults is compared to the class all children.

The point is that constructions which are potentially ambiguous as to the nature of the comparison (collective or distributive) are interpreted as distributive unless marked to the contrary.

Superordinate class and subclass can be compared collectively. For some of these comparisons, e.g.: numerosity (the class-inclusion task), weight and physical extension, the result is known apriori to anyone who understands the nature of the comparison and the class relations, e.g.

(20) Are there more children than just the boys?

(21) Which weigh more: all the coins in my pocket or just the pennies?

(22) Which is larger: our block or the whole city?

Hence in the real world, requests for these specific comparisons must occur infrequently. (Notice that all are syntactically marked.) Other collective comparisons, those based upon usefulness, preference, value, etc. are more common. For instance

(23) The girls alone do a better job than all the children together.

(24) What do you want: the whole newspaper or just the sports section?

Note again that if the collective comparison markings are omitted, e.g.,

(25) Girls do a better job than children.

(26) What do you want: the newspaper or the sports section?

we tend to reinterpret this as a distributive comparison of coordinate classes e.g. big girls and little children, or we assume the speaker is talking nonsense, making a feeble joke, e.g. sports are not news.

One collective comparison, numerosity, cannot be interpreted as a distributive comparison. However, numerical comparisons of class and subclass are usually marked as collective comparisons even though such markings are redundant. We say

(27) There are more children all together than just the boys.

rather than

(28) The children are more than the boys,

and we say

(29) Which is more: all the animals or just the dogs?

rather than

(30) Which is more: the animals or the dogs?

Thus, contrary to standard usage, the class-inclusion question in its usual form is not marked as a collective comparison.

3. Summary and Hypothesis.

Let us now summarize the previous discussion and formulate an account of children's erroneous performance in the class-inclusion task. First, distributive comparisons of classes are linguistically more basic than collective comparisons of classes, as seen by the fact that collective comparisons are the linguistically marked cases. From this we predict that over-generalizations-if they occur-are from the more basic (unmarked) distributive comparisons to collective comparisons. Second, distributive comparisons of class and subclass are ungrammatical while collective comparisons of class and subclass are grammatical. We hypothesize that young children over-generalize this constraint on comparison of class and subclass from distributive comparisons to collective comparisons; hence collective comparison of class and subclass will be ungrammatical for these subjects. Finally, distributive comparisons are usually of coordinate classes in some hierarchy. We hypothesize that the tendency to compare coordinate classes is also over-generalized from distributive to collective comparisons; hence these young subjects will erroneously compare coordinate classes in the class-inclusion task. These effects are intensified by the fact that the class-inclusion question has an anamolous form, the usual redundant markings that indicate that the numerical comparison is a collective comparison are omitted. Basically we hypothesize that erroneous class-inclusion performance is due to over-generalization of a) grammatical constraints and b) expectations of comparison of coordinate classes.

In the next section (B) we first (Section 1) review psycholinguistic data relevant to over-generalization. Then (Section 2) we will examine another aspect of young children's thinking, lack of analytic aptitude toward language, which we will hypothesize also determines performance in the class-inclusion task.

B. Psycholinguistic Considerations

1. Overgeneralization of rules by children.

It is consistently found in developmental psycholinguistics that rules descriptive of the speech and comprehension of children are over-generalized. Rules which describe a limited set of instances are applied more widely than is appropriate. Over-generalizations found in children's speech (Ervin, 1964; Cazden, 1968) include noun inflections (children say dogs, cats but also feets), and verb-inflections (doed, breaked). Over-generalized rules are also found in children's comprehension. For instance, the order actor-verb-object is taken to apply to both active and passive sentences (Fraser etal, 1963). The boy hit the girl and The boy was hit by the girl are taken as synonomous. In general the over-generalization is from the more common and the more uniform to the rarer, more complicated instance. Hence, in the case of comparisons we expect generalizations from distributive comparisons to collective comparisons rather than vice versa.

Suppose a child over-generalized the restriction on the distributive comparison of superordinate class and subclass so that both distributive and collective comparisons of class and subclass are ungrammatical for him. Such a child could not comprehend correctly the question in the class-inclusion task. Would such a child none-the-less answer some other question? Existing evidence suggests he would.

The relevant evidence concerns the tendency of the listener to interpret ungrammatical material in accord with his own grammar. Natural speech includes very many sentence fragments and syntactic anomalies; yet listeners understand. They impose grammatical structure on speech in the service of comprehend-ability. A number of studies bear on this issue, e.g. parents' systematic expansions into grammatical sentences of children's telegraphic, elliptical speech (Brown and Bellugi, 1964); subjects' reinterpretations of semantically implausible compounds as plausible ones, in spite of the task requirements (Gleitman and Gleitman, 1970).

Further as C. Chomsky (1969) reports of her experimental subjects: "... we find that the children do in fact assign an interpretation to the structures that we present to them. They do not, as they see it, fail to understand our sentences.

They understand them, but they understand them wrongly (p. 2)." This description is applicable also to the class-inclusion "non-conservers", they too understand wrongly.

What wrong understanding is expected in the class-inclusion task? We have argued above (Section 2 of A) that we expect distributive comparisons to be of coordinate classes. If the child over-generalizes this expectation to collective comparisons, then he will compare coordinate classes in the class-inclusion task.

In sum we have argued 1) that distributive comparisons are more "basic" linguistically than collective comparisons, 2) that over-generalization of constraints from the more basic distributive comparisons to collective comparisons a) interferes with the child's ability to make a numerical comparison of part and whole, and b) causes him to erroneously compare coordinate classes, part and part, in the class-inclusion task.

2. Analytic Aptitude

A second factor which appears to influence performance in the class-inclusion task is what we will call analytic aptitude. A collection of cognitive skills described by a number of different writers all seem to presuppose the ability to analyse verbal material. We mean by analysis the extraction of information, not the reduction of a whole to component parts. Extraction of information from verbal material must involve relating the object of the analysis to the subject's knowledge, to what has been described as the aperceptive mass. The notion of analytic aptitude is vague because of the lack of a precise model of the process of analysis; what we can do is to indicate the kinds of information yielded by analysis by refering to existing systems, e.g. logic, syntactics, semantics.

In some tasks successful performance requires that verbal input be analysed at the logical level, e.g. the ability to evaluate tautaulogies and contradictions (Osherson and Markman, 1973). Tasks that require the child to articulate what is wrong with grammatically devient sentences require analysis at the syntactic level (Gleitman et al, 1973). When a child is able to distinguish referent and meaning or to indicate awareness of the arbitraryness of word and referent he is apparently analysing processes (Osherson and Markman, 1973). Discussion of word meanings, production of synonyms and paraphrases (Cazden, 1972) all require semantic analysis. In all of the tasks noted above, performance improves with age; this improvement we attribute to an increase in analytic ability with age.

The hypothesis that analytic ability improves with age is used to account for two phenomena associated with the Piagetian class-inclusion task. As was mentioned in the discussion of Piaget's analysis of the class-inclusion task (1 of Section A above) a few children achieve correct answers in the class-inclusion task by counting. A difference in analytic aptitude, we hypothesize, distinguishes the child who solves the class-inclusion task by counting the lions and animals from the child who answers correctly by utilizing his knowledge that lions are animals. The child who counts interprets the question correctly; his is not mislead by grammatical constraints, but he lacks the ability to analyze the question so as to utilize his knowledge of the class relations of lions and animals. Piaget claims the Stage II child counts because he lacks knowledge about the implications of class relations, however we are suggesting that the child who counts is merely unable to utilize such knowledge in this task (i.e., unable to take an analytic approach), he is deficient not in knowledge but in the mobilization of his knowledge.

The second phenomena relevant to analytic aptitude is the finding that some adults (Klahr and Wallace 1972) and some older children (as we will later show; see Section D) misinterpret the class-inclusion question initially but, with repeated presentations of the task, they interpret it correctly. Presumably the older person errs for the same reason the child does, over-generalization of grammatical constraints and the expectation that coordinate classes are to be compared. However, by changing his behavior with repeated presentations of the same kind of questions, the older person shows he <u>can</u> extract information from the question which leads to the correct interpretation. This we hypothsize is due to the greater analytic aptitude of the older person.

Two studies in the literature suggest that performance is facilitated, at least for some children, by manipulating the situation so that the child is forced to take a more analytic tack in the class-inclusion task. Smedslund (1964) found that young children (ages four years three months to six years 2 months) did significantly worse if the stimulus material was visible when the class-inclusion question was asked than if the objects were covered. Wohlwill (1968) reports that in two studies class-inclusion performance was better when presentation was completely verbal (e.g. the child was told the number of objects in each subclass) than if physical objects were present when the question was asked. In both cases children do better if they can only contemplate the verbal questions and their stored knowledge; physical supports in the form of concrete objects interfere with correct solutions. If limiting the input to verbal questions leads to a more careful or more complete analysis of the question, then we have accounted for the Wohlwill and Smedslund findings with the notion of analytic ability.

C. Experimental Evaluation of the Grammatical Constraint Hypothesis.

1. Facilitating correct performance.

We hypothesize that children "understand wrongly" in the class-inclusion task because of grammatical constraints: the request to compare class and subclass collectively cannot be understood. That is, we hypothesize that children who can compare class and subclass misunderstood the sentence used to inquire about this knowledge. Thus coordinate classes are compared due to an erroneous generalization from distributive comparisons. If we can predict from this grammatical constraint hypothesis the conditions that should improve the non-conservers performance, then we can test this hypothesis.

We have worked with two factors which we believe may facilitate performance: (a) the class-inclusion question to the subject may be reformulated so as to make its intent more transparent, in accordance with the marking requirements on subordinate-superordinate comparisons; and (b) the nature of the hierarchy formed by the subclasses and the superordinate class may be varied.

More specifically if the over-generalization hypothesis is correct, then correct answers in the class-inclusion task might be elicited from non-conservers if clues are provided to emphasize that the comparison is collective. Performance should improve if the child is asked to compare all the animals with all the dogs. Further help might be provided to the child by emphasizing the asymetry in the class relations, i.e. asking the child to compare all the animals with only the dogs. Note that this is the conventional wording or marking for collective comparisons. As for the nature

of the hierarchy, performance might be facilitated by using classes which are members of preexisting semantic hierarchies (e.g. animals, dogs, and cats; children, boys and girls). Familiarity with unique names for both superordinate class and subclass should facilitate the tendency to treat such classes as wholes. The tendency to treat the class as a whole should be maximized if a) the entire class is designated by a singular term e.g. class (in school), family, team, and b) the class is of limited extension and known membership so the child has treated this class as a whole in other contexts, e.g. his own class in school, his own family, etc..

There appears to be no previous experimental work in which grammatical clues to class relations and to the collective nature of the comparison were varied. There has been previous work in which the nature of the hierarchy was varied: classes have ranged from members of well-established semantic hierarchies (e.g. flowers, animals, children) to classes named by a single attribute (e.g. brown and white beads, red and blue circles). There is no clear picture of the effect of a preexisting semantic hierarchy on performance by children in the class-inclusion task. In one report (Piaget, 1965) class-inclusion performance is best with pictures of animals, intermediate with flowers and poorest with beads; results that "indicate clearly that the use of classes which have specific names is an aid in differentiating between them and forming the hierarchy" (p. 168). However in another report (Inhelder and Piaget, 1964) correct performance is manifest at a later age with semantic hierarchies (specifically ducks and birds and animals) than with single attribute hierarchies (red and yellow primulas, brown and white beads). This opposite result is attributed "to the fact that these classes [animals] are more remote from everyday experience and hence more abstract" (p. 110).[4] As the quotes attest, varying the nature of the stimuli will never provide a test of Piaget's model. Ahr and Youniss (1970) also used classes in a preexisting hierarchy (dogs and cats) as well as classes defined by a single attribute (red and yellow flowers) but do not report on the effects, if any, of this variation. Dodwell (1962) used dolls (boys and girls), tools (rakes and hoes) and cars (yellow and red) and found no difference in performance with the three different sets of stimuli.

2. Underlying processes.

The hypothesis that non-conservers fail in the class-inclusion task because of over-generalization of a grammatical constraint explains _why_ the children make errors but it does not explain _how_. The hypothesis does not specify the underlying processes that mediate the erroneous performance. At least two different kinds of processes could intervene: _misencoding_ or _temporary misdefining_. The former processes would involve "hearing"and/or storing the question with different words than those in the experimenter's utterance, e.g. "dogs" and "cats" would be stored rather than "dogs" and "animals". The latter process would involve temporary redefinition of the superordinate term, e.g. "animals" would have cats as referent. These two possibilities can be distinguished by asking the child to compare dogs and animals while presenting more cats than dogs. If the child misencodes rather than misdefines, he should reply "cats," an alternative not mentioned in the question. If the child misdefines, he should reply "animals" when he is referring to cats as the larger subclass.

Analogous possibilities exist for the Piagetian model; when the whole is destroyed and reduced to the part, the part may be given the name of the whole or else the part may be named. Inhelder and Piaget report investigation of this issue; their subjects often responded with the name of the superordinate class.

The Klahr and Wallace (1972) information processing models, which attribute failure in the class-inclusion task to the inability to count the same item twice, specify that the non-conserver is always attempting to count the two sets named in the class-inclusion question. Thus their models make a specific prediction: if asked to compare dogs and animals when presented with more cats than dogs, the non-conservers should reply "animals."

Experiment I

Implicit in the preceeding discussion are a number of empirical issues. We report here five studies relevant to these issues. The first is a methodological study that establishes that two slightly different forms of the question requesting a numerical comparison of part and whole are equivalent. One form

Are there more lemons or fruit?

is similar to the form used in previous studies (Ahr and Youmiss, 1970; Dodwell, 1962; Smedslund, 1964; Wohlwill, 1968) of class-inclusion with English speaking children. The second form

Which is more: the lemons or the fruit?

is used in subsequent studies reported here because grammatical clues to class relations and collective comparison can be varied more readily with this form.

a. Subjects. Nine children from professional, academic or business, middle-class families who attended suburban public schools served as subjects. Their ages ranged from five years-eleven months to six years-ten months, with a median age of six years, five months.

b. Method. Children were seen individually either at school or at home. An experimental session lasted from 5 to 15 minutes depending upon the child's interest in extraneous conversation. All children appeared to enjoy the sessions.

A trial consisted of placing a set of objects made up of two coordinate classes in front of the child with a rough grouping into the subclasses (e.g. all the lemons on the left).

As the experimenter placed the objects in front of the child she introduced them by naming the superordinate class: "Here is some fruit." She then asked for a numerosity comparison of the superordinate class and one of the two subclasses, e.g. lemons and oranges were presented and the question referred to fruit and oranges.

Questions by the child were evaded with "Um" or "What do you think?" The child's final answer was counted, e.g. if the child responded "the fruit, no, the oranges", oranges was scored as the answer. The experimenter indicated the child's response on a data sheet. Sessions were recorded on tape to obtain exact wordings of question and response. Trials are omitted from analysis if the question was misworded, or if an interruption occurred.

c. Stimuli. Six sets of objects were used: blocks (red and white Lego blocks of the same shape and size), marbles (large and small glass marbles), sticks (long and short wooden dowels), animals (small plastic lions and giraffes), fruit (plastic lemons and oranges), and children (boy and girl dolls). Note that both the superordinate class and the subclass have distinctive names in the latter three sets, e.g. animals, lions and giraffes. These are instances of preexisting semantic hierarchies. In the first three sets of objects the subsets are named by modification, e.g. blocks, red blocks, and white blocks.

The larger subset contained either three, four or five objects; the smaller subset contained one or two less. The exact number of objects, as well as which class was named first, the subclass or the superordinate class, varied at random from trial to trial.

d. Design. Each child was asked twelve questions: six questions of the form

(31) Are there more lemons or fruit?

and six questions of the form

(32) Which is more: the lemons or the fruit?

Each of the six sets of objects were used once with each question form. All questions mentioned the superordinate class and the larger subclass. Questions of the two forms were presented intermixed in random order.

e. Results. Two versions of the question requesting a numerical comparison of part and whole were presented, one form,(31), has been used in previous studies with English speaking children. Neither question form contained syntactic clues to collective comparison nor to the class relations. As can be seen in Table I, performance with the two forms were virtually identical. Thus the form of the question,(32),used in the following experiments appears comparable to the usual Piagetian class-inclusion question. Given these results the subsequent experiments are considered studies of the Piagetian class-inclusion task.

There was no difference in performance with single-attribute hierarchies and with preexisting semantic hierarchies.

Experiment II

In this study we varied grammatical aids to correct comparison in the class-inclusion task. The experimental sequence consisted of four series of questions. It began with a series of six questions (Series I) in which no clues to class relations or collective comparison were given. These trials correspond to the Piagetian class-inclusion task. In the second series (Series II) clues to class relations and collective comparison were provided to determine if more informative and more conventional wording would facilitate correct interpretation of the question. The third series (Series III) were of the original form to determine the carry-over, if any, from the wording of the second series. Finally (Series IV), six questions were asked in which the clues were deceptive.

a. Subjects. Fourteen children who served as subjects (Subjects 1-14) were in the first grade of private or public schools with similar academic programs. Their ages ranged from six years-two months to seven years-six months with a median age of six years, eleven months. Ten children who served as subjects (Subjects 15-24) attended a private school in which grades one, two and three were combined. These children were all in the same class-room and ranged in age from six years-seven months to nine years-six months, with a median age of eight years- four months. The median age of the total group of twenty-four subjects is seven years-one month.

b. Method. Children were seen individually either at school or at home. Experimental sessions lasted from ten to thirty minutes. Otherwise the procedure duplicated that of Experiment I.

c. Stimuli. For subjects 1 through 14 the stimuli were the same as in Experiment I. For Subjects 15 through 24 dishes (plastic cups and plates) were used instead of children (boy and girl doll).

d. Design. Twenty-four questions were asked of each child: the first series of six trials consisted of questions with no clues to collective comparisons and class relations, as in (32) above. In the second series of six trials the questions contained clues to the comparison and to class relations:

(33) Which is more: only the lemons or all the fruit?

The questions in the third series of six trials again contained no clues as in (32). The questions in the final series of six trials contained misleading clues to class relations.

(34) Which is more: all the lemons or only the fruit?

All questions referred to the superordinate class and the larger subclass.

Between each series of questions of the same form the experimenter paused and said "Here are some more questions." If the child asked, as a few did, if the questions were the same, the experimenter replied with a non-commital "Wait and see."

Eight children who were available were seen a second time, a week or more after the first session. They were rerun on the first two series of questions, question forms (31) and (32).

d. Results. The date from previous studies and from Experiment 1 reported here (Table 1) show that children's performance in the class-inclusion task tends to be all-or-none, i.e. correct on most trials or wrong on most trials. These findings make it inappropriate to average data over subjects, rather subjects should be categorized. In the analysis of Experiment II, criteria are needed to classify children on their Series I performance as either being capable or incapable of solving the class-inclusion task. By Piaget's criteria and terminology these are "conservers" and "non-conservers" respectively. In practice it is easy to classify the children: conservers answer correctly four or more times on the first six trials, non-conservers make five or more errors on the first six trials (Table II).[5] Twenty-one of the 24 subjects are not conservers according to their Series I performance. Analysis here is confined to the results for these 21 non-conservers.

In Series II the wording of the question gives clues to the comparison and to class relations (<u>all the fruit or only the lemons</u>). Seven children who were non-conservers with the first series of questions answered Series II questions correctly. The two forms of questions require the same numerical comparison of subclass and superordinate class; the only difference is the presence of additional clues to collective comparison and class relations.

What changes when a change in wording changes performance? Three further results with these seven children are informative. First, consider Series III in which the form of the question is the same as in Series I - the usual class-inclusion question. Five of the seven children consistently responded correctly, although the question is in the form they did not answer correctly when the experiment began. Thus the correct answer for the majority of these children cannot be dictated by the presence of "all" and "only" in the question.[6] (Experiment III below offers additional evidence for this conclusion.)

The second result that is instructive is performance on the fourth group of questions when the use of "all" and "only" is opposed to the class relations. The five children who responded correctly in Series II when "all" and "only" appear in the questions and in Series III without such clues continue to respond correctly when the use of "all" and "only" is opposed to the class relations. These results make it clear that for these children the correct answer is not tied to "all", neither to its presence nor to its previous use.

An additional informative sign of what the change in wording does, or fails to do, for the child comes from children rerun a week or more later. The effect of the clues was apparently gone: the three children who had responded correctly with "all" and "only" forms of the question in the first experimental session did not answer correctly in the rerun when asked questions without clues -- e.g.

"Which is more: fruits or lemons?" When clues to class relations were again provided, two of the three again answered correctly, but one did not. (Four children, who were seldom if ever correct in the original experimental session, persisted in making wrong responses throughout the rerun). We cannot be sure that there was no "saving" in the ability to correctly compare class and subclass for these two children, but it is slight compared to short term (within a single session) effects.

In sum, one-third of the non-conservers we studied were able to respond correctly when the wording of the class-inclusion question provided additional information about class relations and collective comparison.

Further if the introduction of clues to collective comparisons and class relations is effective in eliciting a correct comparison in the class-inclusion task, then correct comparisons tend to be made in the short run even when the clues are omitted or reversed. These clues either <u>set</u> the child for the comparison specified by the question or weaken the determiners of the faulty comparison or do both. Long term effects are not apparent; at least they were not grossly apparent for the children tested. Finally we have evidence from one subject that syntactic clues can lead to correct judgments on some occasions -- but not on others.

Again there was no difference in performance with single-attribute hierarchies and with preexisting semantic hierarchies.

Experiment III

In the preceeding experiment one-third of the non-conserving subjects responded correctly after grammatical clues to the correct comparison were provided. We concluded that the clues led to the correct interpretation of the question. An alternative explanation is possible: the introduction of the clues may have signaled to the child that his previous answers were wrong and that "all" indicates the correct answer. Although correct answers in Series III and IV of the previous study cast doubts on this alternative explanation of the results, a more direct test is provided by this study.

We first asked children to make numerical comparisons of coordinate classes by asking

(35) Which is more: the giraffes or the lions?

and then provided misleading clues to the correct comparison by asking

(36) Which is more: only the lions or all the giraffes?

with more lions than giraffes presented.

a. Subjects. Eight children whose ages ranged from five years eight months to six years nine months served as subjects. All were first grade students in suburban public schools.

b. Method. The procedures were the same as in Experiment I.

c. Stimuli. The stimulus objects were the same as in Experiment I.

d. Design. A series of six questions were asked of the children in which numerical comparison of the two subsets was requested.

The form of the question (35) is given above. Each of the six sets of objects was used once. Then a second series of six questions were asked in which the smaller subset was referred to as "all" ("all the giraffes") and the larger subset was referred to as "only" ("only the lions") as in (36) above.

e. Results. As expected, the children had no difficulty when asked to compare two subsets. There was only one error in forty-eight trials. When the wording of the question was changed so that the smaller subset was referred to as "all the lemons", and the larger subset was referred to as "only the oranges", there were no errors. Thus introduction of "all" and "only" into the question is not sufficient to cause children to change their responses. Such a change in wording leads to a behavior change only in special cases such as the class-inclusion task where the expanded wording provides clues to the correct interpretation (Experiment II).

Experiment IV

Piaget claims that performance in the class-inclusion task provides an index of mastery of hierarchical classification because it requires comparison of part and whole. We have cast doubts on this claim by showing that more conventional and more informative wording of the request to compare elicited correct responses from some children who were consistently wrong in the usual class-inclusion task. However it is possible that Piaget is correct. The children who answered correctly after the change in wording indeed may have mastered hierarchical classification but, for some unspecified reason (unspecified in a Piagetian context), they may have found this class-inclusion task too difficult.

One way to establish that the class-inclusion task does not measure mastery of hierarchical classification as Piaget claims, is to show that class-inclusion performance and the ability to compare part and whole in other ways are independent. To do this we asked some of the children who participated in Experiment II, the class-inclusion study, their preference for part or for whole. For instance, we asked:

> Suppose you were going to give a party, your Mother said you had two choices: to invite the boys or the class, Which would you invite: the boys or the class? Why?

a. Subjects. Ten children who were subjects in Experiment II were asked questions of preference. Seven of these children were consistently incorrect in Experiment II, one child was consistently correct, and two children were correct with clues in Experiment II. Subject 3, but no other child in this study, also participated in the rerun of Experiment II.

b. Method. The children were questioned individually immediately after completing the last trial of Experiment II. The experimenter said she would like to ask some different questions and asked the child if he would like to continue. All agreed.

c. Design Twelve questions were asked of each child. All began with "suppose". Trial 1 is given verbatim above. The next five trials involved deciding which to take on a trip: "your three favorite toys or your toys"; changing either "your shirt or your clothes"; asking "your mother or your family to a party"; using "the red crayon or the crayons"; or inviting "the girls or the class" to a party. Six more requests for preference were then given which involved the same choices as the first six trials (in a different order) with additional grammatical clues to the comparisons: "only the boys or everyone in the class".

d. Stimuli. Physical objects were not used.

e. Results. Table III gives the children's responses when asked which they would choose to ask to a party: "the class or the boys" in one instance, the "class or the girls in the other. Of the ten children who participated, seven gave unambiguous indications that they actually compared part and whole.[7] Six of these children did not respond correctly in the class-inclusion task, not even when helpful clues were provided. The conclusion is clear: children <u>can</u> compare class and subclass, part and whole, who cannot respond correctly in the class-inclusion task. Further, four of the non-conservers in the class-inclusion task (Subjects 3, 8, 9, 13) actually mentioned that the class was "more" than a subclass. Thus, class-inclusion non-conservers can compare class and subclass quantitatively in some other context.

With a few exceptions the children's responses on the other trials were ambiguous; we could not tell if their choice was between the superordinate class and the subclass or between two subclasses.

Two children (subjects 8 and 13) could not, or would not, justify their choices on some trials without clues but consistently did so when clues were presented. Unfortunately we cannot tell if the ability to justify is related to the presence of grammatical clues or merely to additional trials in the experiment.

Experiment V

The final study examined performance for two numerical comparisons: comparison of a class with the smaller of two subclasses, and comparison of a class with the larger of two subclasses - the usual class-inclusion task. Results of this study provide information about what the child who errs is actually doing.

Comparison of the whole and the smaller part permits one to distinguish which of two possible processes mediates erroneous responding in the class-inclusion task: a) misencoding the question as a question about part and part, "lions" and"giraffes", rather than part and whole, "lions" and "animals", or b) redefining the name of the whole, "animal", to refer to a part, the giraffes. Consider the child presented with five giraffes and three lions and asked "which is more: animals or lions?" If the child misencodes he will reply "giraffes"; if the child redefines "animals" he will reply "animals."

a. Subjects. Eight children served as subjects who were first grade students in a suburban public schools. Their ages ranged from five years ten months to seven years seven months, with a median age of six years, three months. These children did not participate in any other sutdy.

Data from one additional/child is omitted from analysis. Preliminary examination of the date revealed that this child named the <u>larger</u> objects when the objects of the two subsets differed in size, otherwise he selected the larger subclass.

Inquiries revealed he had been referred for neurological examination on the basis of class-room performance and psychological testing. Clearly he was an atypical child and his results do not belong with those of the other subjects.

b. Method. The procedure is the same as in Experiments I, II and III.

c. Stimuli. The same sets of objects were used as in Experiment I.

d. Design. Each child was asked to make twelve comparisons. The first six trials requested a comparison of the superordinate class and the smaller subclass, the last six trials requested comparison of the superordinate class and the larger subclass, the usual class-inclusion task. All questions were of form (32):

<u>Which is more: the giraffes or the animals</u>?

e. Results. When asked to compare the superordinate class with the <u>smaller</u> subclass, seven of the eight children responded with the name of the larger subclass at least once; that is, they gave a response which was <u>not</u> one of the alternatives in the question. (Table IV). This we take as direct evidence that children are actually comparing coordinate classes when they fail to compare part and whole correctly. However, in this same task six children responded with the superordinate class at least once, four children did so on a majority of the six trials. In spite of these responses it seems inappropriate to regard the children as Piagetian conservers on the basis of this evidence. On the following six trials, the usual class-inclusion task, all children chose the larger subclass.

Apparently both hypothesized processes, misencoding and redefining, operate in children: some children appear to misencode consistently, Subjects 1 and 8 always name the larger subclass; some children appear to redefine consistently, Subjects 5 and 6 usually name the superordinate class (Table IV).

The data for the other four children deserves more detailed analysis (Table V). This data presents the only sign in this series of experiments of differential performance with two different ways of naming subclasses. With subclasses named by modification these children tend to respond with the names of the larger subclass, they apparently misencode by adding a modifier to the name of the superordinate class. With generic names for subclasses these children tend to respond with the name of the superordinate class, they apparently redefine. Overall, the children tend to aviod changing wording, where appropriate they add rather than change words.

D. Discussion

1. Grammatical constraints. We hypothesized that the young child's failure in the Piagetian class-inclusion task is attributable to inappropriate generalization from distributive comparisons to collective comparisons. The prohibition on the distributive comparison of class and subclass is inappropriately generalized to collective comparisons. We predicted that class-inclusion performance (which involves a collective comparison, numerosity) would be facilitated: a) by providing additional grammatical clues indicating that the requested comparison is collective, and b) by using classes from well-established semantic hierarchies, which have familiar generic names, so that both class and subclass have been treated as wholes by the child.

The effectiveness of grammatical clues was clearly demonstrated: correct responses when clues were present were obtained from one-third of the children who could not respond correctly to the usual class-inclusion question. Further, we have shown that the facilitating effect was not tied to specific wording; correct responses continued in the short run when clues were omitted. However, the facilitating effect was not permanent; correct responses decreased to the level obtained in the usual class-inclusion task when questions without clues were asked a week later. Both these findings are consistent with the view that the grammatical constraint acts as a <u>set</u>, a set that can be weakened or overcome by various manipulations of stimulus and situation, but which reestablished its effect in time.

The second prediction was not confirmed. Classes representing well-established semantic hierarchies did not yield performence superior to that with classes from single-attribute hierarchies. In retrospect, this result is not surprising. Even though the generic classes are familiar, nonetheless, distributive comparisons (Giraffes are taller than lions, Lions are the fiercest animal) are more common than collective comparisons (Lions are more numerous than giraffes, Giraffes are a less endangered species). Since spontaneous numerical comparisons occurred with familiar classes and subclasses the child has previously treated collectively (class in school and boys or girls in the class) as shown in Experiment IV, we now predict that the use of such classes in the class-inclusion task would facilitate performance.

Additional evidence, albiet indirect evidence, for the grammatical constraints hypothesis comes from Experiment IV in which class-inclusion non-conservers were able to compare part and whole. Clearly, the non-conservers' difficulties in the class-inclusion task derives from specifics of the task rather than a general deficit in the ability to compare part and whole. We conclude that inappropriate grammatical constraints account in part for erroneous performance in the class-inclusion task.

2. Analytic Aptitude. A second hypothesized determiner. of performance in the class-inclusion task is the subject's

analytic aptitude. We hypothesized that older children and adults surmount the grammatical constraints that lead to wrong responses because they are better able to attend to, analyze, and utilize the information in the question. The analytic stance evidently requires effort. "This is so hard" reported one of the few children (Experiment I, Subject 1) who consistently responded correctly; while children who erred said "It's so simple" and "These are easy" (Experiment II, Subjects 7 and 16).

All four children(Experiment I, Subject 1; Experiment II, Subjects 1, 11 and 21) who were conservers by Piaget's criterion erred on the first trial in the class-inclusion task; this indicates that they were vunerable to inappropriate grammatical constraints. However these children were able to overcome the effects of the constraints and interpret the question correctly on subsequent trials. (A second error, if it occurred, was within a few trials of the first). The hypothesized analysis of the question is clearly suggested by the children's verbalizations. For instance, after an error on the first trial Subject 21 engaged in the following dialogue:

Trial 2: E: Which is more: the red blocks or the blocks?

S: What?

E: Which is more: the red blocks or the blocks?

S: (laughs) The red blocks or the --- what?

E: Or the blocks.

S: The blocks. (laughs)

Trial 3: E: Which is more: the oranges or the fruit?

S: (laughs) Let's see, oranges or all them together? 'Course, all of them together.

When clues are reversed, this is the only child to comment and he does so immediately.

Trial 19: E: Which is more: all the red blocks or only the blocks?

S: Hm, 'all' or 'only'! It should be 'all the blocks or only the red blocks' but it's all the blocks.

As we mentioned before, children's performance on the class-inclusion task was usually all-or-none; all the answers were correct or all the answers were incorrect on a series of trials. A few children deviated from this pattern in Experiment II, in which clues to the correct interpretation were varied. Subject 15 whose performance is most inconsi tent, revealed in his questions that he was struggling with the notion of comparing class and subclass after exposure to clues.

E: Which is more: the lions or the animals?

S: The animals put together?

E: Which is more: the lemons or the fruit?

S: Which is the fruit, these?

E: Which is more: all the red blocks or only the blocks?

S: You mean all the blocks put together?

Subject 8 was struggling with the interpretation of the question when the clues were present.

E: Which is more: only the girls or all the children?

S: All-do you mean all the children together?

E: Which is more: all the blocks or only the white blocks?

S: I still don't know what you mean by that!

He was able to reply correctly on only one-half of the trials with clues, and on no other trials.

Let us now compare deductive solutions to empirical solutions. To determine how a child arrives at a correct answer, one would have to ask him to justify his answer. We did not do this because preliminary investigation suggested that children would be reluctant to justify what they considered to be sel-evident answers. However, some children did offer spontaneous justifications, including a conserver (Subject 11, Experiment II) who appeared to arrive at correct answers empirically.

> E: Which is more: only the giraffes or all the animals?
>
> S: Let me think. All the animals! I was thinking. You know how I know 'all the animals'? 'Cause I looked at four and two and I knew there was six. I was thinking how old I was and I knew I was six so I knew there was more animals. So I was thinking in my head.

We believe that children who arrive at correct answer empirically are best described as having achieved relative freedom from inappropriate grammatical restraints, while still lacking analytic sophistication. Most verbalizations of the conservers suggest that they are solving the problems deductively. Yet the fact that both empirical and deductive solutions occur suggests to us that grammatical constraints and analytic aptitude independently affect class-inclusion performance.

3. Underlying Processes.

The results of Experiment V, in which class and smaller subclass were compared, indicates that the set for a comparison of coordinate classes can operate in two possible ways: it can cause the child to miscode the names of the classes to be compared or it can cause a temporary redefinition of the superordinate class name to refer to a subclass. On the surface the readiness of some children to misname the larger subclass seems inconsistent with findings that children are resistant to the notion that name and referent are arbitrary. Yet there is a clear difference. Here we have a looseness in the meaning of a specific word as seen by its use; non-arbitrariness has to do with verbalizations about the necessary relationship between word and referent.

We find that some children consistently misencode, other children consistently misname. A few children evidently utilize both processes in such a way as to change overt wording as little as possible: "blocks"is misencoded as "red blocks" but "animals" are redefined as lions .

4. Other Explanations

Let us briefly consider other explanations of class-inclusion performance. The Klahr and Wallace (1972) information processing models assume that the child attempts to count the superordinate class in the class-inclusion task but actually counts the elements in the non-named subclass. Hence these models predict the child would consistently respond with the name of the superordinate class when the non-named subclass was larger. This does not happen.

For Piaget, performance in the class-inclusion task is a sufficient index of mastery of hierarchical classification. Our results question the validity of this task as an index of mastery of anything. Failure can be changed to success by providing grammatical clues to the correct interpretation of the experimental question - essentially by providing more conventional wording. Further, Piaget's interpretation of class-inclusion performance is difficult to reconcile with our finding that children who err in the class-inclusion task can make spontaneous numerical comparisons of class and subclass (Experiment IV).

Inhelder and Piaget (1964) argue that a child who can state that "dogs are animals" does not necessarily realize that the class of dogs is included in the class of animals; animalness may be just another property, such as four-footedness. However in some circumstances we should accept the child's statement as a sign of knowledge about class-inclusion; the confusing class-inclusion task may be just such a circumstance. Some of the children who consistently erred in the class-inclusion task reported spontaneously that the subclass was included in the superordinate class. They would reply "lions" when asked to compare lions and animals numerically; and then they would volunteer "Lions are animals, too." We suspect they knew what they are talking about.

E. Conclusions

We have argued that, contrary to Piaget's position, the class-inclusion task is not an adequate index of mastery of hierarchical classification. Empirical support for this position comes from two of the studies reported here. In one study children's class-inclusion performance varied with changes in the wording of the question with meaning constant. In another study children who could not compare part and whole in the class-inclusion task were able to compare part and whole when a preference was requested -- they even volunteered numerical justifications. However to question the use of the class-inclusion task as a sign of mastery of hierarchical classification is not to question Piaget's formulation of hierarchical classification. It remains a most comprehensive characterization of an important cognitive activity and deserves more direct examination.

We have argued that the young child's tendency to over-generalize regularities in the linguistic domaine accounts for failure to answer correctly and for the specific errors in the class-inclusion task. This hypothesis is supported by the finding that some children's class-inclusion performance is vulnerable in a predictable way to changes in the wording of the class-inclusion question.

An additional factor, variation in analytic aptitude, was hypothesized to account a) for children who respond correctly in the class-inclusion task by counting the elements in the whole and in the part, as well as b) to account for the disappearance of errors in the performance of children who are predominately correct in their responses. This ill-defined notion deserves more explicit characterization; such a formulation probably presupposes a more explicit model of the process of comprehension.

Footnotes

1. Here and elsewhere we will refer to children who respond correctly in the class-inclusion task as "conservers" and those who err as "non-conservers" i.e. we give the terms a limited operational meaning. One purpose of this paper is to determine whether or not correct performance in the class-inclusion task is diagnostic of a state or stage in logical cognitive development.

2. Bever and Rosenham (1971) do not specify whether the grammatical restrictions they discuss are syntactic or semantic. In the case of restrictions on comparisons one might argue that such restrictions are what N. Chomsky (1964) has called selectional rules.

3. There is no reason to doubt that these restrictions also apply to French. We asked a native French speaker for an English paraphrase of the French version of "Boys are faster than children" and she provided "Older boys are faster than young children". After brief discussion of the issue she provided a number of instances of comparison of class and subclass in which the French and English versions were equally incomprehensible, e.g. "Oranges are sweeter than fruit".

4. Kohnstamm (1967) has questioned the latter conclusion on the grounds that children do not consider birds to be animals.

5. Subject 1, a conserver by our criteria of performance on Series I questions, exhibits a mysterious pattern of responses when subsequent trials are considered. She is neither a consistent conserver nor a consistent non-conserver, however she is completely consistent in that a) for subclass with generic names she always selects the same subclass (e.g. lions) in Series III and IV whether lions form the larger subclass or the smaller subclass, and b) for subclass named by modification she names the larger subclass when it is short sticks, little marbles or white blocks, otherwise she names the superordinate class. The complete consistency of her behavior indicates that she interpreted the task as something quite different from the class-inclusion task-at least after the first series of questions.

6. Probably for Subject 14 the modifiers all and only do determine performance. He was one of three subjects (the other two never answered correctly) who always included "all" in their answers in Series IV e.g. "All the red blocks".

7. Of the other three children, two gave ambiguous data, and one gave puzzling data. The ambiguous data came from two children (Subjects 11 and 12) who say they prefer guests of their own sex and select "the class" over the unpleasant opposite sex. (Does "class" mean girls or both boys and girls when Subject 11 selects it?). The puzzling data comes from a child (Subject 10) who expresses a distaste for girls but prefers to invite the class rather than the boys to a party. Did he really misspeak?

Table I

Experiment I. Numerical comparison of part and whole with two different wordings of the request to compare: "Are there more lions or animals?" and "Which is more: the lions or the animals?". Cell entries are the number of responses of each type.

		Form of the Question					
		"Are there more _____?"			"Which is more_____?"		
Response:		Whole	Larger Part	Smaller Part	Whole	Larger Part	Smaller Part
Subject*	Age						
1	6-5	5	1	0	5	1	0
2	5-11	0	6	0	0	6	0
3	6-2	1	4	1	1	4	1
4	6-2	0	6	0	0	6	0
5	6-7	0	6	0	0	6	0
6	6-10	0	6	0	0	6	0
7	6-8	4	2	0	0	6	0
8	6-4	1	5	0	0	6	0
9	6-7	0	6	0	1	5	0

* Subject number indicates the order in which the children were seen.

Table II

Experiment II. Numerical comparison of part and whole as a function of grammatical information in the request to compare. All questions began "Which is more:_____?" and were completed in one of three ways: Series I and III by merely naming the classes, ("lions or animals"); Series II by providing valid clues to class relations, ("only the lions or all the animals"); and Series IV by providing deceptive clues to class relations ("all the lions or only the animals"). The children are grouped on the basis of their performance on Series I and II of the original testing. Cell entries are the number of correct responses in a series of six questions. Because of experimenter error or interruptions fewer than six trials were scored in some series. Cell entries with stars are adjusted values. The data for eight children who never made a correct response is omitted, however their ages were used in calculation of median age.

Table II

Series		Original Test				Retest	
		I No Clues	II Clues	III No Clues	IV Misleading	I No Clues	II Clues
Subject	Age						
Consistently Correct							
1	6:10	5	0	3	1	5	1
11	6:11	4	5	5	6		
21	9:0	4	6	6*	6		
Median Age 6:11							
Correct With Clue							
2	6:7	0	6	5	6	0	0
12	7:6	0	6	5	6		
14	6:6	0	6	0	0		
15	7:5	1	5	2	3:6*	0	6
17	8:2	0	6	5	6		
18	6:7	0	5	6	6	0	6
19	8:6	0	5	5	6		
Median Age 7:6							
Consistently Incorrect							
5	7:1	0	1	1	0	0	0
7	6:5	1	0	0	0		
8	6:2	0	3	0	0		
10	6:10	1	1	0	0		
[illegible]	8:3	1	0*	0	0*		
23	8:7	0	2	0	0		
Median Age 7:2							

Table III

Experiment III. Preference comparison of part and whole. Children's responses when asked which they would invite to a party and why. The choices were A "the boys or the class" and B "the girls or the class." Subjects are grouped by their responses on the class-inclusion task. Starred answers are responses to the second presentation of the question i.e. with clues to class relations for children who responded "Don't know" to the first presentation.

Table III

Comparison Requested

			A: "boys or the class"		B: "girls or the class"	
Subject	Age	Sex	Choice	Reason	Choice	Reason
C-I: Consistently Correct						
11	6:11	F	Class	Don't like boys	Girls	I like girls
C-I: Correct With Clues						
12	7:6	M	Boys	Don't like girls	Class	Last time I did "the boys"
14	6:6	M	Class	Not fair, other people in class would stay home	Class	Not fair, girls hav fun, boys stay hom
C-I: Consistently Incorrect						
3	6:10	F	Class	It's bigger, it's more children	Class	To invite boys to party
6	6:5	M	Class	Not fair to girls	Class	I like some boys too
7	6:5	F	Class	Wouldn't be nice just to invite boys	Class	You should like everybody
8	6:2	F	Class	*There'd be more people	Class	There'd be more people
9	7:0	F	Class	Cause they have more	Girls	Don't like boys
10	6:10	M	Class	They're my friends	Class	Girls bother me
13	7:3	M	Class	*To have more people	Class	*More people to play with

Table IV

Experiment IV. Numerical comparison of part and whole when the question names 1) the whole and the smaller part or 2) the whole and the larger part. All questions were of the form "Which is more: the lions or the animals?" Cell entries are the number of responses of each type.

Comparison Requested

		Whole and Smaller Part			Whole and Larger Part		
Response:		Whole	Larger Part	Smaller Part	Whole	Larger Part	Smaller Park
Subject	Age						
1	6-8	0	6	0	0	6	0
2	6-2	3	2	1	1	5	0
3	5-11	4	1	1	0	6	0
4	7-7	5	1	0	1	5	0
5	6-4	6	0	0	0	6	0
6	5-3	4	2	0	0	6	0
7	6-2	1	5	0	0	6	0
8	5-10	0	6	0	0	6	0

Table V

Numberical comparison of class and smaller subclass for two ways of naming subclasses: generically and by modification. Data is presented for the four children who responded differentially to the two ways of naming on the first six trials of Experiment II. A comparison of the whole and the smaller part was requested. Cell entries are the number of responses of each type.

Method of Naming Subclasses

	Generic Naming			Modification Naming		
Response:	Whole	Larger Part	Smaller Part	Whole	Larger Part	Smaller Part
Subject						
2	3	0	0	0	2	1
3	3	0	0	1	1	1
4	3	0	0	2	1	0
7	1	2	0	0	3	0

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

Bever, T. G. & Rosenbaum, P. S. Some lexical structures and their empirical validity. In Semantics (Danny D. Steinberg and Leon A. Jakobovits, Eds.). London: Cambridge University Press, 1971. Pp. 586-599.

Birkhoff, G. & MacLane, S. A Survey of Modern Algebra. New York: Macmillan, 1958. Pp. 335-341.

Brown, R. & Bellugi, U. Three processes in the child's acquisition of syntax. Harvard Educational Review, (1964) Vol. 34, No. 2, 133-151.

Cazden, Courtney B. The acquisition of noun and verb inflections. Child Development, 1968, 39, 433-448.

Chomsky, Carol. The acquisition of syntax in children from 5 to 10. [Research Monograph Series No. 57] Cambridge, Mass.: M.I.T. Press, 1969.

Ervin, S. M. Imitation and structural change in children's language. In E. H. Lenneberg (Ed.) New directions in the study of language. Cambridge, Mass.: M.I.T. Press, 1964.

Fraser, C.; Bellugi, U.; & Brown, R. Control of grammar in imitation, comprehension and production. Journal of Verbal Learning and Verbal Behavior, 1963, 2, 121-135.

Gleitman, L.& Gleitman, H. Phrase and paraphrase, New York: W. W. Norton, 1970.

Gleitman, Lila R.; Gleitman Henry; & Shipley, Elizabeth. The emergence of the child as grammarian. Cognition, Spring 1973, 1 (2/3).

Hayes, John R. The child's conception of the experiments in Information Processing in Children. Sylvia Farham-Diggory, (Ed.) New York: Academic Press, 1972.

Inhelder, B. & Piaget, J. The early growth of logic in the the child. New York: W. W. Norton, 1964.

Kessel, Frank S. The r[illegible] of syntax in children's comprehension from ages six to twelve. Mono. Society for Research in Child Development. 1970, 35, (6), 95.

Klahr, D. & Wallace, J. G. Class inclusion processes. In Information Processing in Children, Sylvia Farnham-Doggory (Ed.). New York: Academic Press, 1972.

Kohnstamm, G. A. Piaget's analysis of class inclusion: right or wrong? The Hague: Mouton, 1967.

Osherson, Daniel & Markman, Ellen. Language and the ability to evaluate contradictions and tautologies. University of Pennsylvania, 1973. Unpublished manuscript.

Piaget, J. The child's conception of number. New York: W. W. Norton, 1965.

Podwell, P. C. Relation between the understanding of the logic of classes and of cardinal number in children. Canadian Journal of Psychology, 1962, 16, 152-160.

Swedsland, J. Concrete reasoning: a study of intellectual development. Monog. of the Society for Research in Child Development, 1964, Serial No. 93.

Vygotsky, L. S. Thought and Language. Cambridge, Mass.: M.I.T. Press, 1962.

Wohlwill, J. F. Responses to class inclusion questions for verbally and pictorially presented items. Child Development, 1968, 39, 449-466.