Journal of Experimental Child Psychology **74,** 69–106 (1999)
Article ID jecp.1999.2510, available online at http://www.idealibrary.com on IDEAL®

Developmental Sequences in Class Reasoning and Propositional Reasoning

Ulrich Müller, Bryan Sokol, and Willis F. Overton

Temple University

This study explored the theoretical prediction that class and propositional reasoning skills emerge as a function of the developing ability to coordinate increasingly complex negation and affirmation operations. Children from Grades 1, 3, 5, and 7 (7-, 9-, 11-, and 13-year-olds) were presented with problems from each domain. Rasch analyses of the children's responses were consistent with the hypothesis that both types of problems measured a single underlying dimension (i.e., the coordination of affirmation and negation operations). Qualitatively distinct levels of class and propositional reasoning were identified along this dimension, adding support to the notion that children's reasoning follows a logical developmental sequence. Planned comparisons supported the order-theoretical prediction that different groups of items account for solution differences between grade levels. Results also indicated that children encounter significant difficulties when they have to reason on the basis of negative information.

Key Words: class reasoning; propositional reasoning; developmental sequence; negation; affirmation; Piaget; Rasch scaling.

The identification of behavioral developmental sequences is a central task of developmental inquiry. Developmental sequences indicate that cognitive abilities are acquired in an ordered and invariant manner and that certain levels of cognitive development precede, and are a necessary condition for, the emergence of succeeding levels of cognitive development (Campbell & Richie, 1983; Overton, 1998; Piaget, 1955/1977c; Schröder, Edelstein, & Hoppe-Graff, 1991). Importantly, behavioral developmental sequences are distinguished from contingent, nonnecessary orders by virtue of their being grounded in a theoretical perspective that makes specific order-theoretical predictions about how children

The authors thank the principals, teachers, and students from the participating schools for their cooperation; Donna Gilles for assistance in data collection; and three anonymous reviewers for very helpful comments and suggestions. Portions of this paper were presented at the Biennial Meeting of the Society for Research in Child Development, Washington, DC, April 1997.

Address correspondence and reprint requests to Ulrich Müller, now at the Center for Adoption Research and Policy, University of Massachusetts, 55 Lake Avenue North, Worcester, MA 01655; to Bryan Sokol, now at the University of British Columbia, Department of Psychology, 2136 West Mall, Vancouver, BC V6T 1Z4, Canada; or to Willis F. Overton, Temple University, Department of Psychology, Weiss Hall, Philadelphia, PA 19122. E-mail: Ulrich.Mueller@banyan.ummed.edu, Bsokol@interchange.ubc.ca, or Overton@vm.temple.edu.

0022-0965/99 $30.00

will perform on tasks that tap the cognitive abilities of interest. As succinctly expressed by Schröder et al. (1991, p. 167), “the concept of developmental sequence commands explanatory status only if theoretical and conceptual arguments support the postulated order of acquisition.”

In this study we examine a set of developmental sequences that has been theorized to exist in the domains of children’s class and propositional reasoning, and we explore empirically the proposed developmental relationship between them. The various tasks selected to test these proposed sequences, as well as the order-theoretical predictions associated with how children will perform on them, are largely derived from Piagetian action theory. Before we go into the details of these predictions, it is thus necessary to explicate how the predictions are grounded in the broader context of this theory.

According to the “relational perspective” presented in Piaget’s developmental action theory, the emergence of new knowledge is explained by the construction of increasingly complex relationships between meanings (Piaget, 1968/1971; Piaget & Garcia, 1991; see Overton, 1998). This constructive enterprise can be seen particularly clearly in children’s developing abilities to coordinate logical operations (Piaget, 1968/1971; see also Müller & Overton, 1998; Müller, Sokol, & Overton, 1998) and to develop, in turn, larger operational knowledge structures that can be described as being governed by specific rules of combinations (Piaget, 1947/1950, 1968/1971). Central to this entire process is what Piaget (1974/1978, 1974/1980a, 1975/1985) has termed the coordination of the operations of affirmation and negation.

From this perspective, every cognitive action, whether a sensorimotor action or a mental operation, entails both affirmative and negative components. During the sensorimotor and preoperational periods, children are not cognizant of the negative aspects of their cognitive acts. During these developmental phases, attention is centered on the positive or affirmative aspects of overt actions and mental operations. The initial primacy of affirmations over negations is rooted in the fact that perception focuses on observable, positive qualities, and action pursues positive goals. In contrast, negative qualities cannot be directly perceived, but must be constructed inferentially.

Because of these difficulties in constructing negations, it is not until the phase of concrete-operational thought and the construction of operational structures that each affirmation explicitly has a corresponding negation (Piaget, 1974/1980a). The coordination between affirmation and negation “consists of bringing a reverse operation, and thus a negation, into correspondence with *each* direct operation or affirmation” (Piaget, 1974/1980a, p. 297). Through this coordination, negation becomes an intrinsic variation within a broader system of transformations, and this is a prerequisite for the reversibility of thought and deductive reasoning (Piaget, 1974/1980a, 1981/1987; see Overton, 1990; Smith, 1993).

The coordination of affirmation and negation also transforms the reference frame within which the meaning and extension of concepts are determined. To illustrate, take the statement “This is not a dog.” This statement “does not determine under which concept the object belongs, but solely that it belongs in the sphere outside A [the concept]" (Kant, 1800/1974, p. 110). The positive determination of the object requires the detachment of the negation from the predicate and its coordination with an affirmation; otherwise the meaning of the object would remain vacuous (Piaget & Garcia, 1991, pp. 73–74).

Despite its importance in Piaget's later writings (Piaget, 1974/1978, 1974/1980a, 1975/1985), the idea that reasoning develops as a function of the coordination of affirmation and negation has not received much attention in recent developmental theorizing (see as exceptions Kesselring, 1981, 1985; Riesz & Cantor, 1985). This proves to be particularly unfortunate, since the idea offers a clear empirical opportunity to test the assertion that developmental states unfold to their own intrinsic logic (Kesselring, 1993), characterized at first by a relative lack of differentiation and later becoming more differentiated and hierarchically integrated (Overton, 1998; Piaget, 1975/1985; Werner, 1957). For Piaget, in particular, this developmental understanding was expressed in terms of the thesis that genesis emanates from a structure and culminates in another structure (Piaget, 1967, p. 147). In the case of reasoning development, this is simply because every determination implicitly entails a negation (Piaget & Garcia, 1991, p. 3), and by becoming reflectively aware of this negation—often elicited through contradictions, or oppositions, from others (Piaget, 1965/1995)—children learn to differentiate the negation from and coordinate it with the affirmation.

Piaget used children's class reasoning as his prime example to demonstrate how the progression to higher levels of operational thought reflected such increasingly complex coordinations of affirmation and negation (Inhelder & Piaget, 1959/1969; Piaget, 1977a, 1975/1985). Specifically, his theory asserts that children's ability to solve increasingly difficult class reasoning problems reflects their developing skill at coordinating affirmations and negations. In the present study, various class reasoning problems including subclass–subclass comparisons, class inclusion, double complementation, vicariant inclusion, the law of duality, and, finally, logical implication were used to assess a set of Piagetian-based predictions.

Class Reasoning

Subclass comparison and class inclusion. The inclusion relation between hierarchically related classes is basic to class reasoning. A typical class inclusion task asks children to compare the number of objects in the including or superordinate class with the number of objects in the most numerous of two of its subclasses. For example, given 12 daisies and 4 roses, children are asked, "Are there more daisies or more flowers?" A correct answer requires that children conserve the including class (B) while making the quantitative comparison between it and the included class (A).

Although this may sound simple enough, such a comparison actually involves a multistep process in which children must not only be able to construct the including class, but also be able to reverse this affirmative operation by properly decomposing it. The first step involves being able to combine two subclasses to form a superordinate class, or $A + A' = B$ (Piaget called A the *primary* subclass and A′ the *secondary* subclass).

The second step involves performing the inverse (negative) operation associated with this combination of subclasses. This entails subtracting each subclass from the superordinate class such that $A = B - A'$ and $A' = B - A$ (Piaget, 1977a; Smith, 1982). The inverse operation, thus, implies that children construct each subclass through negation under the including class (Inhelder & Piaget, 1959/1969; Piaget, 1941/1952b, 1975/1985). Piaget termed this type of negation *partial* because it is applied to a part of a larger whole. Through partial negation children realize that the subclass A is an autonomous whole, which enables them to recognize that there are some B's that are not A's (e.g., there are some flowers that are not daisies) and that, therefore, there are more B's than A's.

It has been suggested that preoperational children fail to answer inclusion questions correctly because they only understand negation as a simple form of *otherness* (Inhelder & Piaget, 1959/1969). This simple comprehension of negation leads to splitting those elements belonging to the primary subclass A and all of B's other elements. Failure in class inclusion tasks indicates that the negation has not yet become an intrinsic variation within a larger system of transformations and that the meaning of the subclass has not yet been relativized with respect to the superordinate class. As a result, the integrating whole (B) is no longer conserved when it and a primary subclass (A) are compared. Rather, due to the dominance of affirmations, the whole is identified with one of its subclasses (Piaget, 1941/1952b, 1977a).

Subclass comparisons (e.g., daisies vs roses) by themselves, unlike class inclusion questions, involve only positive, observable characteristics. For this reason, they should be the easiest sort of class reasoning problem for children to solve.

In the present study, we were also interested in exploring the relation between class inclusion problems with wide and restricted frame of reference. Inhelder and Piaget (1959/1969, pp. 109–110) mention in passing that one child in their study solved a class inclusion question with a frame of reference that was restricted to the experimental condition but was unable to solve a class inclusion question with a wide frame of reference (e.g., "*In the woods*, are there more primulas or more flowers?"). Because the developmental relation between class inclusion tasks with restricted and widened frames of reference has not been studied in any systematic fashion, we made no specific predictions regarding their relation.

Vicariant inclusion and double complementation. Closely related to the understanding of hierarchically related classes and the emergence of partial negations is another form of classificatory thinking, which Piaget termed *vicariance operations* (Piaget's concrete-operational grouping II; Piaget, 1972, 1977b; see Chapman, 1988; Davidson, 1987; Flavell, 1963). Two characteristics are basic to vicariance operations. First, the union of a subclass with its complement is equivalent to the union of any other subclass with its complement ($A_1 + A_1' = A_2 + A_2' = B$). For example, the class "flowers" consists of roses and nonroses, or daisies and nondaisies, or tulips and nontulips, and so forth. In other words, while the class B can be divided in different ways, these divisions leave the class B invariant. The different divisions form vicariances, or complementary substitutions, of the same whole. The acquisition of vicariant operations thus indicates that thinking has become more flexible and mobile.

The second characteristic of vicariance operations is that they permit the understanding that each primary class (A_1) is included in the complement of another primary class (A_2'). This feature of vicariant operations will be referred to as *vicariant inclusion.* For example, the primary class of roses (A_1) is included in the secondary class of flowers that are not daisies (A_2').

Available data suggest that vicariant operations are relatively late-emerging concrete-operational abilities (Beilin & Lust, 1975; Piaget, 1980b; Riesz & Cantor, 1985). Unlike standard class inclusion tasks, vicariant inclusions still pose major difficulties for most third graders. The only available empirical study of vicariant inclusion to date (Riesz & Cantor, 1985) found that without training less than 20% of all third graders make correct vicariant inclusion judgments. Further, even following extensive training, only approximately 10% of all fifth graders are able to justify their correct vicariant inclusion judgments.

This difficulty with vicariant inclusion problems is most likely due to the fact that success requires children to coordinate and hierarchically integrate partial negations that are constructed on *two* different levels: the level of the superordinate class and the level of subclasses (as opposed to one level in the case of class inclusion). That is, vicariant inclusions require that the secondary class (A'), which has been formed through negation ($A' = B - A$), must be assigned the rank of a superordinate class, as well as be conserved while being compared to a primary class (A_2). Assigning the secondary class the rank of a superordinate class requires a positive determination of the secondary class $A_1' = A_2 + A_3$ (e.g., nonrose = primulas + daisies), while the comparison of the secondary class with a primary class requires the inferential constructions $A_2 = A_1' - A_3$ and $A_3 = A_1' - A_2$. In contrast to class inclusions, where the superordinate class is already supplied, vicariant inclusion problems require the inferential construction of the superordinate class, as well as the construction of the various subclasses. Children's inability to coordinate and integrate the inferential construction of the superordinate class with the inferential construction of its subclasses is reflected in their failure to conserve the secondary class A_1' while making the comparison between primary and secondary class (Riesz & Cantor, 1985).

Clearly vicariant inclusion problems should be more difficult than class inclusion problems. In addition, they should also be more difficult than the comparison of two secondary classes constructed through complementary substitutions. To illustrate, take a set of 12 flowers, consisting of 5 roses, 3 primulas, and 4 daisies. In order to correctly compare the number of elements in the two secondary, or complementary, classes of nonroses and nondaisies, children must realize that one set can be divided into different secondary classes. Such a double complementation problem thus involves the inferential constructions of $A_1' = B - A_1$ and $A_2' = B - A_2$. In contrast to vicariant inclusion, however, these partial negations do not have to be hierarchically integrated with each other. For this reason, vicariant inclusion problems are expected to be more difficult than double complementation problems. On the other hand, because double complementation problems involve a more flexible construction of secondary classes through complementary divisions of the same whole, they should be more difficult than simple class inclusion problems.

The law of duality. Class inclusion, double complementation, or vicariant inclusion does *not* require broadening the scope of negation beyond a single superordinate class. That is, at the level of concrete operations, where each of these instances of class reasoning is understood to fall (Piaget, 1977a), negation of a subclass A ("roses") remains restricted to its complement A′ ("other flowers") under the superordinate class B ("flowers"). Piaget (1977a) has suggested, however, that at the level of formal operations the scope of negation is broadened as negation operations are incorporated into and mediated through a wider network of relations. This can be seen particularly in children's understanding of the law of duality (Piaget, 1977a, pp. 89, 112).

The law of duality states that the complement A′ of a subclass (A) is larger than the complement (B') of its superordinate class (B) (Inhelder & Piaget, 1959/1969). For example, the class of nonroses is larger than the class of nonflowers. Without broadening the scope of negation, the complement A′ is constructed through partial negation of A under B, and so, children are unable to properly quantify A′ in relation to B′ (i.e., they fail to see that nonroses is a larger class than nonflowers). On the other hand, by broadening the scope of negation, the complement no longer consists only of A′ but includes classes under B' as well (e.g., plants that are not flowers).

The broadening of the scope of negation requires that children are capable of moving flexibly in a nested tree of classification (Byrnes, 1988). A class inclusion relation at a given level ($A + A' = B$) must be nested into a class inclusion relation at a higher level ($B + B' = C$), thereby forming inclusions on inclusions (Byrnes, 1988; Piaget, 1977a). Within such a nested tree of classification, the meaning of the complement A becomes relative to a larger structured whole ($C - A$) and can then be related to the complement of its superordinate class ($C - B$). Thus, by making the complement to A (roses) relative to the superordinate class C (plants), the complement no longer consists only of A′ (other flowers), but also includes all classes under B′ (e.g., oak trees). As with the class inclusion task, we were again interested in exploring the possible effect of the frame of reference on performance. For this reason, we included law of duality questions with a wide and with a restricted frame of reference, although again no specific predictions were made.

Logical implication. Closely related to children's comprehension of the law of duality is their understanding of connective (Reichenbach, 1947), or logical

(Piaget, 1977a), implication. Piaget, for instance, claimed (1977a, p. 122) that children master logical implication at the same stage at which they also readily understand the law of duality. This relationship is seen to follow from the fact that the broadening of the scope of negations, which forms the basis of children's understanding of the law of duality, also appears to form the underpinnings of children's competence in tasks involving the logical implications of a particular conditional rule. Four standard syllogistic arguments, in particular, can be constructed around such a rule (O'Brien, 1987; Overton, 1990). Each consists of a major premise (i.e., the conditional or "If . . . then . . ." rule), a minor premise, and a conclusion. In Genevan studies of logical implication, for example, children were presented with the rule that "all watches that were made in September are broken" (Morf, 1957; Piaget, 1977a) and asked to provide the correct conclusion to each of these four argument forms, as follows:

1. *Modus ponens:* "All watches that were made in September are broken; this is a watch that was made in September. Is this watch broken?" The correct answer to this question is that the watch must be broken.
2. *Denial of antecedent:* "All watches that were made in September are broken; this is a watch that was not made in September. Is this watch broken?" The correct answer to this question is "cannot tell" because the conclusion is indeterminate.
3. *Affirmation of consequent:* "All watches that were made in September are broken; this watch is broken. Was this watch made in September?" The correct answer to this question is "cannot tell" because the conclusion is indeterminate.
4. *Modus tollens:* "All watches that were made in September are broken; this watch is not broken. Was this watch made in September?" The correct answer to this question is "no."

Morf (1957) and Piaget (1977a) found that children who were already at the concrete-operational level as indicated by their success in class inclusion tasks often gave incorrect answers to the affirmation of consequent and denial of antecedent arguments, that is, the argument forms with indeterminate conclusions (see also Haars & Mason, 1986). For example, they answered that a watch that is broken must have been made in September and that a watch that was not made in September cannot be broken. Children, thus, neither realized that the superordinate class of broken watches (B) included both watches made in September (A) and watches made in other months (A′) nor realized that watches made in months other than September could be either broken or not broken. In other words, children rendered the conditional rule into a symmetrical, or reciprocal, rule: "All watches that were made in September are broken, and all watches that are broken were made in September." In fact, such a symmetrical interpretation of conditional rules is frequently observed in studies of conditional reasoning (e.g., Barouillet & Lecas, 1998; Girotto, Blaye, Farioli, 1989; O'Brien & Overton, 1980, 1982).

Piaget (1977a) attributed the symmetrical interpretation of logical implication to two different sources, both of which involved children's abilities to perform certain class reasoning operations. First, children fail to construct the subclass A′ under B (i.e., watches that are made in September and that are broken) because the logical implication problem does not provide concrete perceptual support for the construction of the subclass A′. Rather, in contrast to the standard class inclusion task described earlier, this task requires children to inferentially construct the subclass A′ on a purely verbal or propositional plane. Second, children's failure to realize that watches made in months other than September can be either broken or not broken is due to constraints in the construction of negation. The truth of the conditional rule "All watches that are made in September are broken" is compatible with the following three cases: AB, $-AB$ (watches that are not made in September are broken), or $-A-B$ (watches that are not made in September are not broken) (Piaget, 1977a; Smith, 1993). As noted above, children at the concrete-operational level can construct only partial negations under the nearest superordinate class ($-AB$). Due to constraints on negation, then, such children cannot construct the complete complement to AB and, as a consequence, fail to properly see all the logical implications that follow from the conditional rule (Byrnes, 1988; Piaget, 1977a).

Class reasoning predictions. To summarize, Piagetian theory claims that several forms of class reasoning reflect increasingly complex coordinations between affirmations and negations. This coordination process, in turn, leads to the construction of structured wholes that yield increasingly rich meaning implications, which is accomplished through the broadening of the scope of negation and closely related to children's transition to formal-operational thought. Although these forms of class reasoning have been examined separately, the hypothesis that they reflect a single underlying dimension having to do with the coordination of affirmations and negations has never been subjected to an empirical test. That is, the order-theoretical predictions that can be seen to grow out of a Piagetian developmental approach to understanding reasoning have not been directly explored in any empirical manner.

The first purpose of the present study, then, is to examine critically several Piagetian based predictions concerning the hierarchical arrangement of these forms of class reasoning. Specifically, we predicted the following ascending order of difficulty for the class reasoning items: (1) subclass comparisons, (2) class inclusion, (3) double complementation, (4) vicariant inclusion, and (5) the law of duality and logical implication. Following Piaget (1977a, p. 122), it is predicted that the law of duality and logical implication operate at the same level of difficulty, as both involve the broadening of the scope of negation.

A second purpose of this study focuses more directly on formal-operational reasoning. Piaget argued that the broadening of the scope of negation, entailed in the law of duality and logical implication, is only one characteristic of formal-operational reasoning. Piaget never clarified how this characteristic of formal-operational reasoning is related to what he considered its most important feature: the subordination of reality to possibility, or what is generally understood as hypothetico-deductive reasoning (Inhelder & Piaget, 1955/1958). In order to explore this relationship we also administered a propositional reasoning task that required, at its most difficult level, reasoning in a hypothetico-deductive manner.

Propositional Reasoning

The relationship with class reasoning. Previous research by Noelting, Coudé, Fiset, and Brunel (1993) has suggested that children's propositional reasoning competence is a function of their ability to manage increasingly difficult negation operations. The purpose of Noelting et al.'s work, in particular, was to demonstrate how children's ability to coordinate increasingly complex logical operations involving negations and referential displacements could be seen to follow a distinct developmental path, corresponding to Piagetian stages of preoperational, concrete-operational, and formal-operational thinking (Inhelder & Piaget, 1955/1958). The hypothesis explored by Noelting et al. closely follows our own concerning class reasoning. Specifically, their experimental paradigm permits the systematic study of the ability to construct increasingly complex coordinations between affirmations and negations. Moreover, because Noelting et al.'s measure is sensitive to different levels of thought, it lends itself to exploring the developmental relationship between propositional reasoning and class reasoning. For these reasons, we used a revised version of Noelting et al.'s task in this study.

Negation and referential displacement. The general measurement strategy developed by Noelting et al. (1993) takes the form of a series of logical problems, or items, using propositions that provide the needed information to locate an object in one of three hiding places. Participants' reasoning ability is assessed by their performance in solving these items, which become more difficult as the number of transformations contained in the propositions of each item increases. The transformations come in two forms: displacement and negation operations.

The displacement operation, rather straightforwardly, consists of shifting the reference from one location to another (e.g., the object is *there*). The negation operations, on the other hand, are more complicated and form two types, applying to either the propositional or the metapropositional level of the item. First, at the propositional level, a negation operation can bear on the predicate term of the statement. In this case, the negation is actually embedded inside the proposition (e.g., the object is *not* here). Second, at the metapropositional level, negation can apply to the overall truth value of a statement; that is, the negation is used as a metalinguistic predicate (e.g., "the object is here" is *false*) (Englebretsen, 1981).

Using these transformations in differing combinations, Noelting et al. (1993) identified levels of reasoning consistent with Piagetian stages (Inhelder & Piaget, 1955/1958). At the preoperational level, children are capable of applying just one transformation to a single proposition (e.g., the object is *not* here). They do not exhibit the competence necessary to apply one transformation to another transformation (e.g., falsifying the negation: "the object is *not* here" is *false*), nor do they grasp the link between two propositions with the same meaning (e.g., the statement at location A says, “The object is here,” and the statement at location B says, “The object is in location A”).

At the concrete-operational level children exhibit the ability to coordinate two or more transformations. This can be seen in an example of an item that involves three transformations. In this example, the statement shown at hiding place A states, “The object is not here,” and the child is told that this statement is false; the statement at hiding place B states, “The object is not at A,” and the child is told that this statement is true. Children at the level of concrete operations readily grasp, in this example, that the object must be in hiding place B.

Finally, at the formal-operational level, children exhibit an ability to reason in a hypothetico-deductive fashion. For items that require this level of reasoning, children are told how many statements are true and how many are false, but they are not told directly which statement these truth values apply to. To illustrate, the statement at hiding place A says, “The object is not at hiding place B,” and the statement at hiding place B says, “The object is here.” In this case, children are only told that one statement is true and the other is false but that it is unknown which one is true and which one is false.

The solution of this problem requires reasoning from a hypothetical position. That is, first the child must realize that this problem entails two possibilities because either the statement at A or the statement at B can be true. Second, the child must reason through both possibilities by assigning truth values to the respective statements and must draw the necessary conclusions from truth value assignments (e.g., if the statement at location A is true, then the statement at location B must be false, which implies that the object is in location A; or, switching truth value assignments, if the statement at location A is false, then the statement at location B must be true, which implies that the object is in location B). Third, the child must integrate both possibilities and determine the appropriate response (e.g., impossible in this case because the information is contradictory). The solution process for formal-operational items thus requires that children break from their dependency on fixed truth values and engage in a hypothetico-deductive line of reasoning that involves the subordination of reality to one of several possibilities.

As can be seen in various examples, Noelting et al.’s task embodies the structure of an exclusive disjunction. Every statement about hiding place A has necessary implications for hiding place B. When the statements about one hiding place involve negations, the implications for the other hiding place must be inferred. In order to draw these implications, both hiding places must be considered simultaneously, as copossibilities (Noelting et al., 1993; Piaget, 1981/1987). In other words, the meaning of a statement at any one hiding place must be relativized to the whole experimental situation; otherwise the meaning implications of statements at one hiding place cannot be properly determined. Because items involving more transformations require the construction of more complex meaning implications, Noelting et al.’s procedure allows for straightforward order-theoretical predictions: items involving more transformations should be more difficult than items involving less transformations, and items requiring hypothetico-deductive reasoning should be more difficult than items which do not. Among formal-operational items, items requiring the generation and coordination of a larger number of possibilities should be more difficult than items requiring the generation and coordination of fewer possibilities.

As noted above, the ability to draw necessary conclusions from mere possibilities marks the onset of hypothetico-deductive thought and is the hallmark of formal-operational reasoning. Inhelder and Piaget (1955/1958) examined the development of hypothetico-deductive reasoning in the context of experimental situations that required the determination of the cause of a particular result from several possible factors. In order to solve such tasks, a person had to realize that each hypothesis is compatible with several single observations and that each observation, in turn, is compatible with several competing hypotheses.

Hypothetico-deductive reasoning thus entails a differentiation and coordination between conjunctively linked propositions within a hypothesis *and* disjunctively linked alternative hypotheses (Müller, 1999). As a consequence, at the level of formal operations the truth values of statements are no longer fixed, and they become the subject of systematic investigation. This process is particularly evident in the propositional reasoning problem described earlier.

What is less clear, however, is just how class reasoning problems such as the law of duality and logical implication relate to this hypothetico-deductive reasoning competence. That is, in contrast to the hypothetico-deductive reasoning that is necessary for success on formal-operational items in the propositional reasoning task, neither of the formal-operational types of class reasoning problems seems to require the subordination of reality to possibility. For example, although an understanding of logical implication requires the coordination of conjunctively linked propositions, it does not also require their coordination with disjunctively linked hypotheses. Thus, in addition to examining the order-theoretical predictions associated with Noelting et al.'s (1993) propositional reasoning task, we planned to explore the developmental relationship that has been theorized to exist among the law of duality, logical implication, and hypothetico-deductive reasoning by comparing children's performance on these sorts of problems.

METHOD

Participants

Participants were 20 first (11 boys and 9 girls), third (8 boys and 12 girls), fifth (10 boys and 10 girls), and seventh graders (10 boys and 10 girls), recruited from private and public schools in the Philadelphia area and in Lancaster county. The average ages of the four groups were 7 years 3 months (range 6 years 6 months to 8 years 4 months), 9 years 4 months (8 years 8 months to 10 years 2 months), 11 years (9 years 11 months to 12 years 1 month), and 13 years (12 years 3 months to 13 years 9 months).[1] The sample was primarily middle class and diverse in ethnic composition.

Materials and Procedure

All children were interviewed individually, and the interviews were tape-recorded. Parental consent was obtained for each participant before the interviews were conducted.

Class inclusion. Two tasks were administered to assess children's understanding of class inclusion. The materials used in the first class inclusion task were 15 picture cards depicting two different kinds of fruit (9 apples, 6 bananas). The materials were spread out on a rectangular cloth, with the apples on one side and the bananas on the other.[2]

Children were asked to identify (a) all the bananas, (b) all the apples, and (c) all the fruit. Then the children were asked to answer three questions. The first question concerned a subclass comparison ("Are there more apples or more bananas?"), which served the purpose of disambiguating the interpretation of the following standard class inclusion question.[3] The second question was the standard class inclusion question ("Are there more apples or more fruit?"). Finally, the third question widened the frame of reference beyond the immediate

[1] The overlapping age range for Grades 3 and 5 was due to the fact that 4 fifth graders were relatively young (9 years 11 months to 10 years 2 months). We compared these 4 fifth graders with the remaining 16 in terms of their class and propositional reasoning performance. Whereas there was no difference in class reasoning performance, the younger fifth graders performed worse on the propositional reasoning task (average Rasch person parameters were 1.48 and 3.36, respectively; $t(18) = 1.35$, $p = .19$). When these 4 fifth graders were excluded from the statistical comparisons, the difference between Grades 3 and 5 in terms of overall propositional reasoning performance was only marginally significant, $F(1, 72) = 3.30$, $p = .07$, and the difference between Grades 3 and 5 in their performance on IIB items was no longer significant, $F(1, 72) = 2.22$, $p = .14$.

[2] The number of class inclusion items corresponds roughly to the numbers used in Campbell (1991) and Inhelder, Sinclair, and Bovet (1974/1980). The spatial layout of the items in our experiment follows that of Inhelder and Piaget (1959/1969), who had children sort different items into subclasses and then asked the class inclusion question about spatially segregated items. In contrast to earlier studies (see Winer, 1980), a more recent study (Gold, 1986) suggests that the intermingling of items facilitates class inclusion reasoning by enhancing the salience of the superordinate class, which, in turn, reduces the ambiguity of the standard class inclusion question (Carpendale, McBride, & Chapman, 1996). In our study, however, we selected a different procedure for reducing the ambiguity of the standard class inclusion question (see Footnote 3).

[3] A major criticism of the standard class inclusion task is that the class inclusion question itself (e.g., "Are there more apples or more fruit?") is ambiguously worded (Campbell, 1991; Hodkin, 1981; Shipley, 1979; Richard & Leynet, 1994). That is, the question lacks the normal linguistic markers of a part–whole comparison used in everyday language to better specify a particular class contrast (e.g., *only* the apples, or *all* the fruit). As a consequence, the standard question may be confusing to young children or may be understood as a request for a subclass comparison. Several ways have been proposed to remove this ambiguity and to countermand a subclass–subclass comparison, including better perceptual or linguistic marking of the various classes (Campbell, 1991; Chapman & McBride, 1992; Hodkin, 1981; McGarrigle, Grieve, & Hughes, 1978; Shipley, 1979) and asking for a subclass comparison prior to the standard class inclusion question (Carpendale et al., 1996; Smedslund, 1964; Winer, 1978).

experimental condition ("In the whole world, are there more bananas or more fruit?").

The second class inclusion task employed the same procedure but with two different kinds of birds (seven sparrows, five eagles).

Vicariant inclusion and double complementation. Clay balls of differing sizes and colors were used in a series of questions to assess children's understanding of vicariant inclusion. All together, there were 6 red balls (2 big, 2 medium, 2 small), 2 green balls (both big), 3 yellow balls (1 big, 1 medium, 1 small), and 2 blue balls (1 big, 1 small) that were placed in groups on a rectangular cloth according to their color. The numbers of each kind of ball were selected such that incorrect responses would result from either (a) a comparison of the two primary classes, A_1 vs A_2 (e.g., green vs red), or (b) a comparison of A_1 to the rest of A'_2 without A_1 (e.g., red vs blue and yellow, instead of red vs blue, yellow, and red).

Five questions were asked in total. Two questions required children to make subclass comparisons (e.g., "Are there more red balls or more green balls?"). Two questions assessed children's understanding of vicariant inclusion (e.g., "Are there more balls that are red or more balls that are not green?"). A fifth question required children to compare two complement sets with each other (e.g., "Are there more balls that are not yellow or more balls that are not blue?"). Before the vicariant inclusion and the complementary class questions were asked, children identified the balls that were involved in each comparison (e.g., "Show me the balls that are red. Show me the balls that are not green"). We included the identification of the complementary class of balls that are not green in order to ensure that all children had a basic understanding of complementation. No child had problems with identifying the complementary class.

Law of duality. The same content that was used in the class inclusion tasks was employed in two tasks devised to assess the understanding of the law of duality. The first duality task used the picture cards depicting fruit and several additional cards (e.g., a cherry, a lemon, a tree, a flower). Each child was initially asked to point out (a) all the things that were not fruit and (b) all the things that were not bananas. Comprehension of the law of duality was then assessed through two questions. The first duality question required children to make a direct comparison between the secondary subclass (A') and its complement under the superordinate class (B'): "Are there more things that are not bananas or more things that are not fruit?" The second question made a similar request, but involved a change to a wider and more abstract frame of reference: "In the whole world, are there more things that are not bananas or more things that are not fruit?"

The second duality task involved different kinds of birds, but no picture cards were used, and both duality questions referred to *the whole world* as a frame of reference ("In the whole world, are there more things that are not eagles, or are there more things that are not birds?" "In the whole world, are there more things that are not birds, or are there more things that are not animals?").

Logical implication. Children's understanding of logical implication was assessed through a procedure that was similar to the one used by Morf (1957) and that included all four of the standard syllogistic arguments described in the introduction (i.e., modus ponens, denial of antecedent, affirmation of consequent, modus tollens). Children were told a story about a party taking place at their school to be attended by students from their own and a neighboring school. Before the party begins, all the students from their own school decide to wear blue shirts. The conditional rule to be applied by the child participant, then, was that “all students from my own school are wearing blue shirts.” Children were asked to repeat this rule, and a card with a pictorial representation of the rule on it was placed in front of them to serve as a memory cue.

Before questions about each of the syllogistic arguments were asked, it was emphasized that all students from the child participant’s own school followed the rule, and so, were indeed wearing blue shirts at the party. Next, each question was introduced by asking the children to imagine that they and a friend were walking around during the party when the friend drew a particular conclusion about another student that she saw. For instance, the modus ponens question was presented in the following manner: “First, you see a student from your school but you cannot see the color of her shirt. Your friend says: ’That is a student from our school. Therefore, she must be wearing a blue shirt.’ Is your friend right or wrong?”

The other three syllogistic arguments were introduced similarly, but involved changes in the minor premise (i.e., That is “a student from the other school,” “a student wearing a blue shirt,” or “a student wearing a red shirt”) and the conclusion (i.e., Therefore, “the student cannot be wearing a blue shirt,” “the student must be from our school,” or “the student cannot be from our school,” respectively). The question of whether or not the friend’s reasoning was correct always followed the presentation of these various argument forms.

Propositional reasoning. To assess propositional reasoning, we revised the measure developed by Noelting et al. (1993). The materials included three upright cardboard figures representing children (Peter, Jim, and Ann) playing a simple card game. Each puppet figure was shown holding a single card. Booklets containing the statements made by each player were placed in front of the appropriate figures. In addition, cards indicating whether the players’ statements were true or false were placed next to the booklets as the experimenter read from each. Based on the statements and the information provided by the experimenter, children were asked to determine which of the puppet figures held the winning card (i.e., the card with the star on it). Response options were provided in another booklet. One option, *cannot know,* referred to situations in which the information proved to be contradictory, and so, determining the winning cardholder was impossible.

Each round of the game constituted a different logical problem. Thirteen items of varying difficulty levels were devised by systematically combining displacement and negation operations in the players’ statements (see Appendix). The first 9 items involved only two players (Peter and Jim); the last 4 involved all three. Children were familiarized with the format of the problems during an initial instructional phase. During this phase, the child and the experimenter worked together to solve at least two practice problems, including one problem containing contradictory information and one that required formal-operational thinking. After obtaining a child's response for the formal-operational item, the experimenter demonstrated one technique for solving the item by drawing a diagram to illustrate the possibilities embedded in the item and the consequences of the various truth value assignments. Each possibility was represented by a large box with two compartments, and each compartment represented one player. The experimenter worked with the child first through one and then the other possibility. For each possibility, the experimenter read the statements referring to each player. The consequences of these statements were indicated by either placing or not placing a check mark into the compartments, depending on what the statements said about each player. Next, the truth values were considered and check marks were crossed out or inserted on the basis of truth value assignments. The conclusion that could be drawn from each possibility was written next to each box. Finally, the conclusions from both statements were considered together. The children were encouraged to make their own diagrams using paper and a pencil whenever they felt necessary.

A typical problem was presented in the following manner. First, the experimenter read the information contained in Peter's booklet, and then placed the card indicating whether Peter was telling the truth or lying next to the booklet. Second, the experimenter provided the same information for the other players, Jim and Ann. Third, the experimenter pointed to the booklet containing the response alternatives and asked, "Who has the card with the star on it, Peter, Jim, Ann, or cannot know?" Finally, upon offering a response, children were asked to explain, or justify, their answer. Justifications were tape-recorded for later analysis.

The order of class reasoning and propositional items was counterbalanced across grade levels, with the exception of the logical implication task, which always came last. The wording of class reasoning questions was balanced across participants as well (half mentioned the including term first; the other half mentioned the included term). For the propositional reasoning items, Problems 1 to 9 were randomized to establish an order of presentation. Items 10 to 13 always came last because they involved three puppet figures, as opposed to Items 1 to 9, which involved just two. For both class and propositional reasoning items, half the participants at each grade level received the established order, while the remaining half received the reverse.

Scoring

In addition to making a judgment for each item, children were also required to justify their responses.[4] Success or failure, then, was based on a judgment-plus-

[4] Scoring based on correct judgments alone often leads to an overestimation of children's competencies (Chapman & McBride, 1992; Hodkin, 1987; Thomas & Horton, 1997). In order to reduce the likelihood of ascribing competencies to children who do not possess the target concept but

justification criterion. Tapes of 20 subjects (5 from each grade level) were recoded by a second rater. Each tape contained 4 subclass comparison, class inclusion items, and law of duality items; 2 vicariant inclusion items; 1 double complementation and logical implication item; and 11–13 propositional reasoning items. Interrater reliability (percentage of agreement) for the different items was as follows: (a) subclass comparisons, 100%; (b) class inclusion, 93%; (c) law of duality, 88%; (d) vicariant inclusion, 98%; (e) double complementation, 85%; (f) logical implication, 100%; and (g) propositional reasoning, 93%. Disagreements were resolved by discussion.

Class inclusion. Correct justifications included (a) reference to the fact that all objects belong to the superordinate class ("Bananas and apples are fruit") or that a subclass is only one part of the superordinate class ("Sparrows are just one kind of bird; there are lots of other birds than just sparrows") and (b) reference to the two combined sets as being more than one of the sets ("Sparrows and eagles make more than the sparrows"), often in connection with specific numerosities ("Seven sparrows and five eagles make more than just seven sparrows"). Incorrect justifications included repeated assertions of the correct answer ("Because birds are more than eagles") and nonlogical justifications (e.g., "People like fruit," "I see a lot of birds"). In addition, the justification "There are lots of birds and only few sparrows" was also considered incorrect. We concluded that children giving this justification did not construct operational part–whole relationships between the subclass and the superordinate class. Rather, they considered the superordinate class as an aggregate set (Dean, Chabaud, & Bridges, 1981; Fuson, Lyons, Pergament, Hall, & Kwon, 1988; Wohlwill, 1968) and simply compared two dichotomous classes, without considering the subclass as being a part of the superordinate class. This interpretation receives support from the fact that 17 of 19 occurrences (89%) of this justification were given by children when the "whole world" served as the frame of reference. In 15 of 19 cases (79%) this justification co-occurred with an incorrect judgment for the class inclusion question that referred to just the picture cards (i.e., a restricted frame of reference).

Vicariant inclusion and double complementation. Children's justifications were considered correct when they stated that the primary class is included in the complement of another primary class or when their counting behavior demonstrated an implicit understanding of this fact. Correct justifications for the question regarding the comparison of two complement sets included (a) referring to the count of not-blue and not-yellow balls, (b) referring to the fact that more

---

make correct judgments (false-positive errors), it is crucial that children's justifications are taken into account. Empirical evidence concerning the issue as to whether a "judgment-plus-justification" criterion underestimates children's class reasoning competencies (false-negative errors) is not entirely consistent (Chapman & McBride, 1992; Thomas & Horton, 1997). In any case, the issue of which criteria should be applied for the attribution of knowledge to children is not an empirical, but an epistemological, question (Smith, 1992, 1999). We accept the venerable epistemological position that a true belief counts as knowledge only if it is justified.

is left when the blue ones are taken away than when the yellow ones are taken away, and (c) referring to the fact that there are more yellow than blue ones and that therefore the not-yellow are less than the not-blue.

Law of duality. Correct justifications fell into the following categories: (a) Children stated that taking away the objects belonging to the complement of the included class leaves more objects than does taking away the objects belonging to the complement of the including class and (b) children listed classes that are part of the complement of the included class but that are not part of the complement of the including class. The following explanation of a seventh grader illustrates this latter type of justification: “There are more things that are not eagles because not-eagles includes all the other birds that wouldn’t be included in the group of things that are not birds.” Often children made correct judgments, but gave incorrect justifications. In addition to simply guessing, the following types of incorrect justifications were encountered: (a) Children simply compared a primary class (e.g., eagles) to its complement (e.g., not-eagles), and (b) children compared a primary class to its complement, which was constructed through partial negation under the including class (“There are more things that are not eagles because there are a lot more birds other than eagles”).

Logical implication. Only children who made correct judgments and gave correct justifications for all four logical implication questions were scored as passing this task. In particular, to be scored correct, a participant had to recognize the indeterminacy of the denial of antecedent and affirmation of consequent argument forms. For example, children had to answer the affirmation of consequent argument by somehow suggesting that their friend could be either right or wrong, because a student wearing a blue shirt could be from either their own or the other school.

Propositional reasoning. Children’s answers were scored as correct if they supplied a correct judgment and gave a correct justification. Only justifications which demonstrated a child’s competence at coordinating transformations at both the propositional and the metapropositional level were considered correct. For example, if a child justified his or her answer to Item 1 by simply saying “cannot know because both Peter and Jim are lying,” the response was scored as incorrect because no evidence was provided to suggest that the child had coordinated the two levels. Because several attempts to administer formal-operational items to first graders failed (the five selected for testing were completely unable to follow their structure), their performance on these items was scored as being incorrect.

RESULTS

Rationale

Starting from the theoretical position that a progressive sequence of competencies entailing the coordination of negations and affirmations operates as a necessary condition for successful class and propositional reasoning, we predicted that problem solutions from each of these areas would fall along a single dimension of hierarchically related difficulty levels. Accordingly, we chose a scaling method of analysis to test this prediction. Specifically, the Rasch model was selected as it provides a more stringent test of unidimensionality than do other scaling models (Bond, 1995a, 1995b; Kingma & Van den Bos, 1989).

Rasch scaling provides both a statistical test for unidimensionality and information about relative item difficulty (Elliott, 1982; Wright & Masters, 1982; Wright & Stone, 1979). The model is based on the notion that only two factors are responsible for a person's success or failure on any particular problem or item: the item's difficulty and the person's ability. The model produces estimates for both of these by calculating item difficulty parameters (logits), which are determined on the basis of the total number of persons who are correct on each item, and person ability parameters, which are determined by the total number of items that each person successfully answers.[5]

The assumptions underlying these estimations are (a) that a more able person will always have a greater likelihood of success than a less able person and (b) that any person will always have a greater likelihood of success on any easier item than on any more difficult item. According to these assumptions, it is expected that items keep their relative difficulty value across all groups of persons and, similarly, that persons keep their relative ability values across all items (Elliott, 1982; Wright, & Masters, 1982; Wright & Stone, 1979). By defining the probability of success as a joint function of persons and items, the Rasch model maintains the "idea of order" (Wright & Stone, 1982, p. 4; see Overton, 1998).

The observed scores are statistically evaluated in terms of the extent to which they deviate from this notion of order. When items measure the same dimension they will fit the Rasch model (Elliott, 1982). Consequently, if the difference between observed and expected scores is within statistically acceptable boundaries, the items are considered to measure the same underlying construct. Similarly, the validity of order-theoretical predictions can be evaluated by examining the distribution of items. Items expected to measure the same difficulty level should cluster together and not overlap with items that are designed or expected to measure a different difficulty level. In other words, items measuring different cognitive levels should be separated from each other by a larger gap than items measuring the same cognitive level (Fox & Gray, 1997). Whether the difference between item parameters is statistically significant can be evaluated in terms of the standard errors provided by Rasch analysis. Item parameters are significantly (95% confidence level) different from each other when their difference exceeds the sum of their standard errors; that is, Item Parameter 1 minus Parameter 2 is greater than the standard error for Error Item 1 plus the standard error for Item 2 (Wright & Stone, 1979, p. 95).

In addition to Rasch scaling, traditional methods of data analysis (i.e., planned

[5] Rasch analysis incorporates the "uncertainties of experience" such as measurement error by "expressing the model of how person and item parameters combine to produce observable events as a probability. This leaves room for the uncertainty of experience without abandoning the construction of order" (Wright & Stone, 1979, p. 4).

contrasts) were used to test for differences in performance on both class and propositional reasoning items for adjacent grade levels. For these analyses, we used a conventional alpha level of $p < .05$ or better. No gender or order effects were found in this last part of the analysis.

We also conducted an analysis of the types of errors that children made at the different grade levels, and calculated percentages for the most common errors. Error data can provide rich information about children's reasoning abilities (Piaget, 1952a; Siegler, 1996). In particular, by demonstrating that specific errors are closely related to specific levels of competency, error data can lend support to order-theoretical predictions such as ours.

Class Reasoning

The first Rasch analysis was performed on 80 participants and 16 class reasoning items. Rasch analysis iteratively excludes those items that are solved either by every person or by no person. Similarly, those persons who solve either all items or no items are also excluded. If persons or items are excluded by Rasch analysis, it does not mean that they do not fit the model. It simply means that the analysis cannot provide *exact* parameter estimates for these cases because the items are either too easy or too difficult, and the persons are either too able or not able enough. Instead, *rough* estimates are provided for the excluded items and persons as maximum (perfect) or minimum (zero) scores.

In the case of the class reasoning data, a total of four items and 29 persons were excluded. The four items making up all of the subclass comparisons received minimum scores, while 13 persons (9, 3, and 1 in Grades 1, 3, and 5, respectively) received minimum scores and 16 persons (1, 4, and 11 in Grades 3, 5, and 7, respectively) received maximum scores. Of the 51 individuals included in the analysis, only 1 (a seventh grader) did not fit with the model's predictions, which was due to the fact that this particular participant (Rasch person measure $= 4.01$, infit $t = 2.1$) solved three of the relatively difficult law of duality items but did not solve two easier items (i.e., logical implication and law of duality Fruit 1).

Table 1 shows the item parameter estimates and fit statistics for the class reasoning items.[6] Because all items fell within conventionally acceptable boundaries ($t < 2.0$), the analysis supports the prediction that these different forms of class reasoning items measure one underlying developmental dimension. Table 2 displays the joint distribution of items and persons in terms of logits (i.e.,

[6] Rasch analysis produces infit and outfit statistics. Infit is a standardized information-weighted mean square statistic, which is more sensitive to unexpected responses that are nearer to a person's ability level. Outfit is a standardized outlier-sensitive mean square fit statistic that is more sensitive to unexpected responses on items that are more distant from the person's ability level. A person who solves a relatively difficult item but does not solve an easy item produces a *positive* misfit. On the other hand, a *negative* misfit indicates response patterns that are too orderly. Although negative misfit is undesirable from a test-theoretical perspective (Wright & Masters, 1982), it can be expected in developmental studies that deal with qualitatively different levels of cognition. Accordingly, in the present study we will consider only positive misfit in the evaluation of unidimensionality.

TABLE 1
Rasch Scaling for Class Reasoning: Item Statistics

Item	Difficulty estimate	Error	Infit *t*	Outfit *t*	Expected level
Duality Bird 2_a	4.43	.57	−1.6	−0.2	FO
Duality Bird 1_a	4.13	.53	−1.2	−0.2	FO
Duality Fruit 2_a	4.13	.53	−1.7	−0.2	FO
Vicariant Inclusion 1_b	2.01	.39	−0.1	0.2	CO
Logical Implication$_b$	2.01	.39	1.7	0.0	FO
Duality Fruit 1_b	1.86	.39	1.5	0.5	FO
Vicariant Inclusion 2_b	1.71	.39	−0.5	0.1	CO
Double Compl.$_c$	−1.91	.50	0.7	0.6	CO
Incl. Fruit 2_d	−3.36	.58	−0.8	−0.4	CO
Incl. Bird 1_{de}	−4.12	.65	−0.3	−0.2	CO
Incl. Bird 2_{ef}	−5.12	.76	−1.4	−0.2	CO
Incl. Fruit 1_f	−5.79	.86	−0.3	−0.1	CO
Sub. Bird$_g$	−7.56	1.48	Minimum est. measure		PO
Sub. Fruit$_g$	−7.56	1.48	Minimum est. measure		PO
Sub. Vicar. 1_g	−7.56	1.48	Minimum est. measure		PO
Sub. Vicar. 2_g	−7.56	1.48	Minimum est. measure		PO
Mean	.00	.55	−0.3	0.0	
SD	3.65	.15	1.1	0.3	

Note. n = 80 persons, 16 items. Analyzed: 51 persons, 12 items. PO = Preoperational item; CO = concrete-operational item; FO = formal-operational item. Items with different subscripts differ significantly from each other.

difficulty parameters). Items are centered around .00 logits, with higher items being more difficult.

On the basis of the item distribution, five distinct difficulty levels, which were largely consistent with our order-theoretical predictions, were evident. The least difficult level consisted of the subclass comparisons. Not surprisingly, these items were solved by all children. Although somewhat spread out in their difficulty, the next higher level consisted of the class inclusion items. Among these items, those having fruit as their content added support to Inhelder and Piaget's (1959/1969) anecdotal evidence that class inclusion is significantly more difficult with a wide frame of reference (Fruit 2) than with a restricted frame. This difference, however, did not reach statistically significant levels for items having birds as their content. The third level contained only the double complementation item.

The fourth level was made up of the vicariant inclusion and logical implication items, as well as one of the law of duality items for which there was concrete support present (i.e., a picture card). Finally, the fifth level, consisting of the most difficult items, included the duality items for which there was a wide frame of reference and no concrete support (i.e., the picture cards were absent).

Composite scores for the four class inclusion, two vicariant inclusion, and four

TABLE 2
Rasch Scaling for Class Reasoning: Joint Distribution of Persons and Items

Person measure			Item measure
6.0	$n = 16$; 1 3rd, 4 5th, 11 7th	FORMAL-OPERATIONAL	6.0
5.0	$n = 3$; 2 5th, 1 7th		5.0
		Duality Bird 2	
4.0	$n = 2$; 1 5th, 1 7th	Duality Bird 1, Duality Fruit 2	4.0
	$n = 1$; 1 3rd		
3.0		CONCRETE-OPERATIONAL IIC	3.0
	$n = 10$; 2 1st, 4 5th, 4 7th		
2.0		Duality Fruit 1, Log. Impl., Vicariant Inclusion 1	2.0
	$n = 8$; 1 1st, 2 3rd, 4 5th, 1 7th		
		Vicariant Inclusion 2	
1.0			1.0
	$n = 7$; 5 3rd, 2 7th		
0.0			0.0
	$n = 7$; 1 1st, 4 3rd, 2 5th		
−1.0		CONCRETE-OPERATIONAL IIB	−1.0
−2.0		Double Complementation	−2.0
	$n = 6$; 3 1st, 1 3rd, 2 5th		
−3.0		CONCRETE-OPERATIONAL IIA	−3.0
		Inclusion Fruit 2	
	$n = 3$; 2 1st, 1 3rd		
−4.0		Inclusion Bird 1	−4.0
	$n = 2$; 1 1st, 1 3rd		
−5.0		Inclusion Bird 2	−5.0
	$n = 2$; 1 1st, 1 3rd	Inclusion Fruit 1	
		PREOPERATIONAL	
−6.0	$n = 13$; 9 1st, 3 3rd, 1 5th	Subclass Comparisons	−6.0

Note. $n = 80$ persons, 16 items. Analyzed: $n = 51$ persons, 12 items. IIA = Easy concrete-operational item; IIB = medium concrete-operational item; IIC = difficult concrete-operational item.

law of duality problems were constructed by totaling the correct responses that children gave to each of these groups of items. The proportions of correct responses to all of the class reasoning problems are presented in Table 3 according to grade level and problem type. For the composite scored items,

TABLE 3
Percentage Correct for Class Reasoning Tasks by Grade Level

Grade	CI	DC	VI	CR	LD
1	.45	.25	.08	.10	.00
3	.73	.65	.15	.20	.13
5	.95	.80	.65	.55	.35
7	1.0	1.0	.83	.80	.65
Total	.78	.68	.43	.41	.28

Note. $n = 20$ in each group. CI = class inclusion; DC = double complementation; VI = vicariant inclusion; CR = conditional reasoning; LD = law of duality.

subsequent planned comparisons were computed for adjacent grade levels with number of correct responses as the dependent variable.[7]

The planned comparisons showed that (a) third graders solved significantly more class inclusion items than first graders, $F(1, 76) = 7.06$, and fifth graders solved significantly more class inclusion items than third graders, $F(1, 76) = 4.73$; (b) fifth graders performed significantly better than third graders on the vicariant inclusion items, $F(1, 76) = 21.41$; and (c) performance on the law of duality items increased from Grade 3 to Grade 5 as well as from Grade 5 to Grade 7, $F(1, 76) = 4.69$, and $F(1, 76) = 9.04$, respectively. For double complementation, two-tailed tests for differences in proportions revealed that third graders performed better than first graders, $z = 2.54$, and seventh graders performed better than fifth graders, $z = 2.11$. Performance on logical implication was significantly better for fifth graders than for third graders, $z = 2.29$.

To test whether there was a general effect for type of reference frame ("the whole world" vs a restricted frame), the two items for each problem type were combined and subjected to a 4 × 2 (grade × problem type) ANOVA with repeated measures on the second factor. The ANOVA revealed significant effects for grade, $F(3, 76) = 11.78$, problem type, $F(1, 76) = 4.75$, and the interaction between grade and problem type, $F(3, 76) = 4.75$. The interaction was followed up with post hoc Scheffé tests. These indicated that first graders performed significantly better on class inclusion problems with a restricted frame of reference than on problems with a wide frame (*M*s 1.0 vs 0.8).

Class Reasoning Error Analysis

Error analysis supported the Piagetian claim that the development of class reasoning consists of the construction of increasingly complex operational structures that instantiate ever more complex coordinations between affirmations and negations. First graders were often incapable of simultaneously coordinating the direct (i.e., affirmative) operation and its inverse in class inclusion tasks. As a

[7] The results were essentially the same when the Rasch person parameter was used as the dependent variable.

consequence, their most common error was the failure to conserve the superordinate class and to simply compare both subclasses (19% of all responses). This error was almost entirely absent in the remaining grade levels.

The double complementation task requires the construction of two partial negations to divide the same whole two different ways. About one third of the first graders failed to construct any negation and simply compared the two primary classes (e.g., the blue and the yellow). This error was rare among third and fifth graders (5% and 10%, respectively) and nonexistent among seventh graders.

Error analysis of performance on the vicariant inclusion items was consistent with the prediction that the difficulty of vicariant inclusion resides in the conservation of a negatively specified superordinate class while comparing it to one of the primary classes. The most frequent error in vicariant inclusion was to ignore the membership of a given primary class (A_1) within the complement of another primary class (A'_2). This error accounted for 68%, 63%, 33%, and 13% of all responses in Grades 1, 3, 5, and 7, respectively.

In order to recognize that the complement of the subclass includes the complement of the superordinate class, children must broaden the scope of negation and construct a nested tree of classification. Inspection of children's errors suggests that third and fifth graders fail to construct such a classification tree. About half of the third graders and one third of the fifth graders simply opposed a positive term (e.g., eagles) and its complement (e.g., not-eagles), which included everything else except the positive term. This kind of reasoning led to the correct judgment for the wrong reasons.

We also observed a failure to construct negations in the logical implication task. For all grade levels, the prevalent error in this task was to construct a symmetry between antecedent and consequent ("All students from our school are wearing blue shirts, all students wearing blue shirts are from our school"), which led to correct responses for the modus ponens and modus tollens arguments but to incorrect responses for arguments requiring indeterminate answers (denial of antecedent, affirmation of consequent). This error, which reflects the failure to construct not-B under A, was committed by 55%, 50%, 35%, and 20% of all first, third, fifth, and seventh graders, respectively.

Propositional Reasoning

Rasch analysis of the propositional reasoning data resulted in the exclusion of one item and 15 persons. The formal-operational item requiring the generation of three possibilities received a maximum score. Six first graders received minimum scores, and 9 persons (1 fifth and 8 seventh graders) received maximum scores. Of the 65 individuals analyzed, only 2 (1 third and 1 fifth grader) did not fit the model. The third grader (Rasch person measure = .44, outfit $t = 2.1$) did not solve the easy Item 9, but solved more difficult items. The fifth grader (Rasch person measure = 1.25; infit $t = 2.0$) did not solve the relatively easy Item 11, but solved the more difficult Item 12.

Rasch analysis supported the primary hypothesis that all propositional reasoning items measured the same construct (*t* values within acceptable boundaries; see Table 4) and, with only a few exceptions, that our order-theoretical predictions regarding item difficulty were correct (see Table 5). More specifically, Rasch analysis supported the predictions that items involving hypothetico-deductive reasoning are more difficult than concrete-operational items and that the item involving the generation of three possibilities (Item 13) is more difficult than the one involving just two (Item 3). Finally, among the concrete-operational and preoperational problems, it was generally the case that items involving more transformations were more difficult than those involving fewer. There were two exceptions to this trend. First, Item 11 involving six transformations clustered together with items involving just two and three transformations. Second, Item 8 involving just two transformations was slightly more difficult than the items involving three transformations.

The number of solved items was used as the dependent variable for a set of planned contrasts with adjacent grade levels as the independent variable.[8] These contrasts revealed that all adjacent grade levels differed significantly from each other: Grade 7 ($M = 10.8$, $SD = 1.80$) compared to Grade 5 ($M = 9.15$, $SD = 1.95$), $F(1, 76) = 5.20$; Grade 5 compared to Grade 3 ($M = 7.25$, $SD = 2.36$), $F(1, 76) = 6.89$; and, finally, Grade 3 compared to Grade 1 ($M = 2.40$, $SD = 2.89$), $F(1, 76) = 44.88$.

Based on the item distribution and the number of transformations contained in each item, five different difficulty levels were distinguished: preoperations, or PO (Item 9); easy concrete operations, or IIA (Items 4, 5, 7, 8, and 11); medium concrete operations, or IIB (Items 1, 2, and 10); difficult concrete operations, or IIC (Items 6 and 12); and formal operations, or FO (Items 3 and 13). Table 6 shows the averaged proportions of solved items for each level by grade.

Subsequent comparisons showed that specific groups of items accounted for grade differences. Success on IIA and IIB items accounted for the higher scores of third as compared to first graders, $F(1, 76) = 96.86$ and $F(1, 76) = 12.38$, respectively. Success on IIB and IIC items accounted for the higher scores of fifth as compared to third graders, $F(1, 76) = 5.50$, and $F(1, 76) = 15.25$, respectively. Finally, success on IIC and formal-operational items accounted for the higher scores of seventh as compared of fifth graders, $F(1, 76) = 4.74$, and $F(1, 76) = 13.01$, respectively.

Propositional Reasoning Error Analysis

Incorrect responses tended to result from a failure to coordinate the propositional and metapropositional levels. That is, children typically failed to integrate their understanding of the statements at the propositional level (e.g., Peter is saying, “I have the card with the star on it”) with the truth values at the metapropositional level (e.g., Peter is lying). Noelting et al. (1993) termed this

[8] The results were essentially the same when the Rasch person parameter was used as the dependent variable.

TABLE 4
Rasch Scaling for Propositional Reasoning Test: Item Statistics

Item	Difficulty estimate	Error	Infit *t*	Outfit *t*	Level
Item 13_a	9.23	1.45	Maximum est. measure		FO
Item 3_a	8.50	1.05	0.0	0.0	FO
Item 12_b	3.06	.40	−0.7	−0.3	IIC
Item 6_b	2.61	.38	1.6	0.1	IIC
Item 10_c	1.11	.37	−0.3	−0.5	IIB
Item 2_c	0.98	.37	−0.6	−0.4	IIB
Item 1_d	−0.18	.40	−0.7	−0.4	IIB
Item 11_e	−1.06	.44	−1.4	−0.4	IIA
Item 8_{ef}	−1.24	.44	0.1	0.0	IIA
Item 4_{fg}	−2.10	.49	0.3	0.1	IIA
Item 7_{fg}	−2.10	.49	0.2	0.0	IIA
Item 5_g	−2.61	.52	0.2	0.0	IIA
Item 9_h	−6.94	1.09	0.4	0.2	PO
Mean	0.00	.54	−0.1	−0.1	
SD	3.62	.24	0.7	0.2	

Note. n = 80 persons, 13 items. Analyzed: 65 persons, 12 items. Items with different subscripts differ significantly from each other. PO = Preoperational item; IIA = easy concrete-operational item; IIB = medium concrete-operational item; IIC = difficult concrete-operational item; FO = Formal-operational item.

type of error *fragmentation.* To illustrate, fragmentation led children to answer Item 2 on the basis of information derived solely from the positive truth value at the metapropositional level. This resulted in the incorrect response of "Peter because he is telling the truth" (30% of all responses to Item 2). In other words, the children centered on the affirmative information in the problem and thus ignored the implications that negating this information might have for the problem's solutions. Fragmentation was a common error across all difficulty levels and generally illustrated children's fundamental problems with coordinating affirmations and negations.

Fragmentation was evident even among the least difficult concrete-operational items (i.e., IIA items). For example, with Item 7, 40% of first graders said that "you cannot know because both are telling the truth." Thus, many first graders simply opposed the affirmative truth values, without being able to coordinate them with the negations on the metapropositional level. This failure to coordinate negations and affirmations between the two levels virtually disappeared among the higher grades' responses to the IIA items (only one fifth grader answered "cannot know" with regard to Item 7), indicating their increasing competence in managing such coordinations.

In their responses to the more difficult IIB and IIC items, however, third graders continued to commit fragmentation errors. For example, the response to Item 2 was often (40%) "Peter because he is telling the truth," and to Item 6

TABLE 5
Rasch Scaling for Propositional Reasoning: Joint Distribution of Persons and Items

Person measure			Item measure
9.0	*n* = 9; 1 5th, 8 7th	FORMAL-OPERATIONAL	9.0
		Item 13	
		Item 3	
8.0			8.0
7.0			7.0
6.0	*n* = 10; 1 3rd, 4 5th, 5 7th		6.0
5.0			5.0
4.0		CONCRETE-OPERATIONAL IIC	4.0
	n = 11; 1 1st, 4 3rd, 4 5th, 2 7th		
3.0		Item 12	3.0
		Item 6	
2.0	*n* = 7; 1 3rd, 4 5th, 2 7th		2.0
	n = 10; 1 1st, 2 3rd, 5 5th, 2 7th		
1.0		CONCRETE-OPERATIONAL IIB	1.0
	n = 7; 1 1st, 5 3rd, 1 7th	Item 6, Item 10	
0.0		Item 1	0.0
	n = 3; 3 3rd		
−1.0	*n* = 2; 2 5th	CONCRETE-OPERATIONAL IIA	−1.0
	n = 4; 1 1st, 3 3rd	Item 8, Item 11	
−2.0		Item 4, Item 7	−2.0
	n = 5; 4 1st, 1 3rd	Item 5	
−3.0			−3.0
	n = 1; 1 1st		
−4.0		PREOPERATIONAL	−4.0
−5.0			−5.0
	n = 5; 5 1st		
−6.0			−6.0
−7.0	*n* = 6; 6 1st	Item 9	−7.0

Note. *n* = 80 persons, 13 items. Analyzed: *n* = 65 persons, 12 items. IIA = Easy concrete-operational items; IIB = medium concrete-operational items; IIC = difficult concrete-operational items.

(55%) "You cannot know because both are lying." In addition, they failed to discover the contradiction entailed in the more complex Item 12. The typical response here was that "Jim must have the card because Ann says so" (45%), showing that they ignored the contradictory information presented by the other players. These kinds of error patterns became less frequent for fifth and seventh graders (25% and less).

TABLE 6
Percentage Correct for Propositional Reasoning Items by Grade Level

Grade	PO	IIA	IIB	IIC	FO
1	.70	.24	.15	.03	—
3	.95	.88	.50	.20	.00
5	1.00	.93	.73	.60	.05
7	1.00	.98	.95	.73	.30
Total	.91	.76	.58	.39	.12

Note. $n = 20$ in each group. PO = Preoperational item; IIA = easy concrete-operational items; IIB = medium concrete-operational items; IIC = difficult concrete-operational items; FO = formal-operational items. The proportions shown are averaged across 1 PO, 5 IIA, 3 IIB, 2 IIC, and 2 FO items.

Despite children's explicitly practicing one formal-operational item in the instructional phase, these items were always the most difficult for the children. Almost all third and fifth graders said that the formal-operational items (3 and 13) were indeterminate because "one has to know who is telling the truth and who is lying." This response indicates that children confused the hierarchical relationship between the metapropositional and propositional levels for these items. Interestingly, this response pattern occurred together with a complete enumeration of all possibilities in five of the fifth graders and in three of the seventh graders. Still, the generation of all such possibilities did not lead them to adequately coordinate the levels of possibility and reality, that is, the metapropositional and propositional levels, respectively. Rather, they still confused these two levels, projecting the former into the latter. Only four children solved the most difficult problem (Item 13). Three of these used paper and pencil to draw a diagram outlining the different possibilities as shown in the instructional phase. All other children failed to make use of paper and pencil for any item.

Conjoint Analysis

To examine the developmental relationship between class and propositional reasoning, all items from both domains were subjected to a Rasch analysis. The conjoined Rasch analysis resulted in the exclusion of five items and 13 persons. As would be expected, the subclass comparison items again received minimum scores, and the most difficult formal-operational item received a maximum score. Five first graders received minimum scores, and 8 persons (1 fifth and 7 seventh graders) received maximum scores. Of the 67 individuals included in the analysis, only 2 (1 third and 1 fifth grader) did not fit the model. The third grader was the same as in the Rasch analysis of the propositional reasoning data. The fifth grader (Rasch person measure = 2.84, infit $t = 2.3$) did not solve the relatively easy propositional reasoning Items 2 and 10, but solved two more difficult law of duality items.

Most importantly, the analysis revealed that all items measure the same underlying construct. As shown in Table 7, the *t* values of all items are within acceptable boundaries. Specifically, the item distribution indicates the following developmental relationship between the class and propositional reasoning items. First, propositional reasoning Item 9, involving only one transformation, is significantly easier than the class reasoning items involving a simple subclass comparison. Second, class inclusion items cluster together with the IIA concrete propositional reasoning items. Third, mastery of double complementation occurs after mastery of IIA concrete propositional reasoning items, but precedes mastery of IIB items. Fourth, logical implication, vicariant inclusion, and the law of duality items without concrete support are more difficult than the IIB, or moderately difficult, propositional reasoning items, and cluster, accordingly, with IIC items. Finally, the law of duality items without concrete support are more difficult than the IIC concrete items of the propositional reasoning task, but far easier than formal-operational propositional items.

DISCUSSION

It has been argued that the assessment of developmental sequences involves three interrelated parts (Overton, 1998): (a) a theory of the developmental processes that are functioning in a given domain, (b) developmentally oriented tasks that require the use of these processes, and (c) statistical tools that appropriately implement the theoretical model under examination.

In the current study, we have followed this three-pronged approach to conducting developing research. As evidenced in our introduction, the postulation of developmental sequences ultimately requires a structural analysis between tasks and operations (Schröder et al., 1991). Such an analysis, however, cannot be conducted without some theoretical grounding. The identification of orderly sequences, in other words, can only take place within the context of a theory of order. Accordingly, we used Piagetian action theory as a theoretical framework for postulating a set of developmental sequences in the domains of children's class and propositional reasoning. Specifically, the idea that development in these domains is a function of the coordination of affirmation and negation offered a context for choosing and constructing specific assessment items. Our second step thus involved developing a set of reasoning problems that required the use of increasingly complex affirmation and negation operations. As this step suggests, item construction is always a theory-guided effort in which the theoretical framework provides a basis for, as well as gives meaning to, the observation of particular behaviors and response patterns (Bond, 1995a; Fox & Gray, 1997; Overton, 1998; Smith, 1991). Finally, we employed Rasch analysis to examine our order-theoretical predictions empirically. The Rasch model's assumptions of orderliness make it both a developmentally sensitive measurement tool (Bond, 1995a, 1995b; Overton, 1998) and congruent with the Piagetian principles that guided this study.

The success of our order-theoretical predictions generally suggests that the

TABLE 7
Rasch Scaling for Conjoint Analysis of Class Reasoning and Propositional Reasoning: Item Statistics

Item	Difficulty estimate	Error	Infit t	Outfit t	Level
Item 13_a	8.64	1.45	Maximum est. measure		FO
Item 3_a	7.05	.80	−0.4	−0.1	FO
Duality Bird 2_b	3.66	.43	−2.3	−0.3	FO
Duality Bird 1_b	3.48	.42	−2.1	−0.3	FO
Duality Fruit 2_b	3.48	.42	−2.4	−0.3	FO
Item 12_c	2.42	.37	1.4	−0.1	IIC
Item 6_{cd}	2.02	.36	1.4	0.4	IIC
Vicariant Incl. 1_{cd}	1.90	.36	0.2	0.3	IIC
Logical Implic.$_{cd}$	1.90	.36	1.0	−0.2	IIC
Duality Fruit 1_{cd}	1.77	.35	0.2	−0.1	IIC
Vicariant Inclusion 2_d	1.65	.35	−0.6	−0.4	IIC
Item 10_e	0.64	.36	−0.5	−0.5	IIB
Item 2_e	0.51	.36	0.1	0.3	IIB
Item 1_f	−0.66	.41	−0.4	−0.5	IIB
Double Complem.$_{fg}$	−1.18	.42	0.0	−0.2	IIB
Item 11_{gh}	−1.55	.44	−1.0	−0.2	IIA
Item 8_{ghi}	−1.74	.44	0.9	0.1	IIA
Incl. Fruit 2_{hij}	−2.14	.45	−2.0	0.5	IIA
Item 4_{ij}	−2.56	.47	0.9	0.2	IIA
Item 7_{ij}	−2.56	.47	1.1	0.2	IIA
Incl. Bird 1_{ij}	−2.56	.47	0.0	−0.3	IIA
Item 5_{jk}	−3.01	.48	0.7	0.0	IIA
Incl. Bird 2_{jk}	−3.01	.48	−0.6	−0.3	IIA
Incl. Fruit 1_k	−3.25	.50	−1.1	−0.3	IIA
Item 9_l	−6.29	.83	1.2	0.2	PO
Sub. Fruit$_l$	−7.96	1.46	Minimum est. measure		PO
Sub. Bird$_l$	−7.96	1.46	Minimum est. measure		PO
Sub. Vicar. 1	−7.96	1.46	Minimum est. measure		PO
Sub. Vicar. 2_l	−7.96	1.46	Minimum est. measure		PO
Mean	0.00	.45	−0.2	−0.1	
SD	2.96	.12	1.2	0.3	

Note. n = 80 persons, 29 items. Analyzed: 67 persons, 24 items. Items with different subscripts differ significantly from each other. PO = Preoperational item; IIA = easy concrete operational item; IIB = medium concrete operational item; IIC = difficult concrete operational item; FO = formal operational item.

coordination of affirmation and negation is an important developmental process that demands more attention. In the following we consider some more specific implications of our results for a developmental account of children's reasoning and, in doing so, show the empirical support that our three-pronged approach to the development of reasoning received.

Class Reasoning

The analysis of the class reasoning items demonstrates that the different forms of class reasoning develop along a single developmental dimension. The ordering of these forms in terms of increasing difficulty is (a) subclass comparison, (b) class inclusion, (c) double complementation, (d) vicariant inclusion and logical implication, and (e) law of duality. With the exception of logical implication, this ordering supports the Piagetian theoretical position that the development of class reasoning entails the construction of increasingly complex operational structures that instantiate ever more complex coordinations between affirmations and negations.

Error analyses and comparison of different grade levels demonstrated that specific class reasoning difficulties tend to arise for each age group. First graders frequently had difficulty simultaneously coordinating affirmation and negation or constructing two partial negations, as required by class inclusion and double complementation tasks, respectively. On the other hand, most third graders failed to conserve a negatively specified superordinate class while comparing it to one of the primary classes, as required by vicariant inclusion. Finally, most fifth graders were still unable to broaden the scope of negation, as required by the law of duality.

The class inclusion questions with a restricted frame of reference were also more readily solved than those with a wide frame of reference. It is unlikely that this effect is due to counting-based solution strategies because, although children referred to the specific counts of the superordinate and subordinate classes to explain their judgments, they were not observed to count before answering the class inclusion question (see also Campbell, 1991). However, the effect of frame of reference emerged only for first graders and was limited to the specific content fruit. Whether the reference frame effect is limited to specific age groups and types of content requires further investigation.

The reference frame effect was clearer among the law of duality items. With concrete support and a restricted frame, understanding the law of duality is easier than without such support and a widened frame. This is most likely because children can employ a counting strategy (which we observed) when picture cards or other support is present.

Although, according to Piagetian theory, both understanding logical implication and understanding the law of duality are based on the broadening of the scope of negation, we found that logical implication is significantly less demanding than the law of duality. It is likely that the specific content of the logical implication problem allowed children to retrieve counterexamples from experience (e.g., children from the other school wearing a blue shirt), which, in turn, made it possible to solve the indeterminate argument forms (see Markovits, 1993; Markovits, Fleury, Quinn, & Venet, 1998; Markovits et al., 1996). If this is the case, then children do not need to simultaneously construct all the possible cases that are consistent with the rule, and they do not need to broaden the scope of negation as they must do in order to solve the law of duality. Rather, for the indeterminate argument forms, they only have to be capable of (a) constructing different divisions of the domain in question (the class *children from the other school* includes both children wearing blue shirts and children wearing shirts of any other color; the class *children wearing a blue shirt* includes both children from our school and children from the other school) and (b) coordinating the primary and secondary class in each case (e.g., $B = A$ or A'; Piaget & Garcia, 1991, p. 28). The ability to construct complementary substitutions of the same domain is a characteristic of vicariant operations, which may explain why logical implication clustered together with these items instead of the more difficult law of duality items. To explore the issue of whether a complete understanding of the intension of the rule is necessary to solve the logical implication task, one could use picture cards and have the participants construct a model of the logical relationships between the classes (see Piaget, 1977a, p. 104—subject Nov).

Propositional Reasoning

The Rasch analysis of the propositional reasoning task supports the prediction that the various items measure different reasoning competencies along a single developmental dimension. The reasoning competencies range from preoperational (Item 9) to formal-operational thinking (Items 3 and 13). Performance on the different groups of items improves by grade level, suggesting that the competencies emerge in an ordered sequence that is age-related. Performance on IIA items increased significantly between Grades 1 and 3. The solution of these items demands that both hiding places must be considered as copossibilities (Noelting et al., 1993; Piaget, 1981/1987). In other words, the meaning of a statement at any one hiding place must be relativized in reference to the whole experimental situation. First graders appeared to have difficulty with coordinating the propositional and metapropositional levels on IIA items. Frequently they simply ignored the negations and reasoned on the basis of affirmations. Fifth graders' performance was significantly better than third graders' on IIC items. These items can be characterized as containing multiple negations, providing no positive information, and demanding the detection of complex contradictions between statements. Other studies have similarly observed children younger than 10 years old struggling to reason with double negation and negative information (Jou, 1988; Piaget & Garcia, 1991, pp. 20–29).

Children at all grade levels encountered major difficulties with formal-operational items, which require coordinating the levels of possibility and reality. From each of the different possibilities, which are disjunctively linked, children must deduce their specific implications, that is, the conjunctively linked cases that would be true if the specific hypothesis were true. Instead of keeping these levels distinct, children tended to collapse the two. A similar failure to distinguish between these levels has also been observed in a study of hypothesis testing (Moshman, 1979). Our study suggests that the differentiation between and the hierarchical coordination of these levels begins to emerge around Grade 7. Future studies using the propositional reasoning measure should extend the age range in order to further examine the development of hypothetico-deductive reasoning.

Although, as expected, formal-operational items were more difficult than concrete-operational items, the order-theoretical predictions for the propositional reasoning problems did not receive complete support. The distribution of the items suggests that certain factors beyond just the number of transformations influence their difficulty level. Specifically, Item 11 involving six transformations clustered, rather surprisingly, with IIA items containing just two or three transformations. Most likely, this is the case because the solution does not actually require performing all six transformations. Instead, Peter's statement may provide an initial anchor around which children can form their judgment using fewer transformations. An interesting study for the future could explore this issue by systematically varying the position of such an anchor.

A second factor that is likely to influence the item difficulty is the particular relationship between propositional and metapropositional levels. Items 5 and 8 are particularly illustrative of this point. Although both items involve two transformations, Item 8 was significantly more difficult than Item 5. It may be argued that these items differ in how their statements at the propositional level and their truth values at the metapropositional level form either a corresponding or a noncorresponding relationship. That is, Item 5 has corresponding statements ("I do not have the card with the star on it") and truth values ("true") for Peter and Jim, while Item 8 does not (the statements are "I have the card" vs "I do not have the card," and the truth values are "false" and "true"). One could argue, then, that the solution of items with corresponding propositional and metapropositional levels requires the construction of less complex meaning implications than does the solution of the items with noncorresponding levels.

Item 5, for example, can be solved by sequential processing: children can exclude first Peter and then Jim and then coordinate the results of their processing. When statements and truth values do not correspond, however, children need to construct meaning implications between the statements of the players. Item 8, for instance, would require the following meaning implications: "If Peter says he has the card with star, and he is lying, then Jim must have the card. Jim says that he does not have the card with the star, and he is telling the truth. Therefore, Peter must have the card with the star." Future studies need to examine in greater detail how the particular relationship between statements at the propositional level and truth values at the metapropositional level influences the construction of meaning implications.

Finally, future research needs to offer the additional response alternative *impossible* in conjunction with *cannot know*. These response options denote very different concepts. Whereas *cannot know* denotes undecidability, *impossible* denotes contradiction. Children in this study used *cannot know* in both ways, and only by taking their justifications into account was it possible to decide the cognitive basis of their response.

Conjoint Analysis

The conjoint analysis of propositional reasoning and the various types of class reasoning shows that the items from both tasks develop along one underlying dimension. In addition, there are systematic relationships between items from both tasks. Again, Item 9 of the propositional reasoning task is much easier than all other items, with the exception of subclass comparisons. The likely reason for the finding that class inclusion items cluster with propositional IIA items is that both require hierarchical reasoning and the simultaneous coordination of at least two operations. Similarly, double complementation, vicariance inclusion, and logical implication (at least as measured here) are structurally similar to propositional IIB and IIC items in that they both require a more complex integration of negations and affirmations.

Although the law of duality is significantly more difficult than the late concrete propositional items, it is less difficult than the formal propositional items. Hypothetico-deductive reasoning and the law of duality both require second-degree operations: hypothetico-deductive reasoning requires the construction of different truth value assignments, and law of duality requires the construction of iterative inclusions (Byrnes, 1988). In contrast to the law of duality, however, hypothetico-deductive reasoning also includes the differentiation of and coordination between various possibilities and reality. In other words, because truth values are no longer fixed in the problems requiring such reasoning, they themselves become the subject of investigation. Prerequisites for formal-operational reasoning, then, are the abilities to generate all possible hypotheses, to work out the implications of each hypothesis for the propositional level, and to coordinate the results of different hypotheses. The finding that even children who are capable of generating all possible truth value assignments may still fail to work out their implications points to the critical role that the differentiation and hierarchical integration of the levels of possibility and reality plays in achieving formal-operational thought. The reasoning suggested by these children's responses highlights their struggle to come to terms with the kind of paradox that formal-operational thought introduces. That is, it poses the problem of how "a conclusion may be necessary at the same time that its truth value depends on the truth value of the hypothesis" (Piéraut-le Bonniec, 1980, p. 100). Future studies are needed to examine children's developing abilities to structure reality on the basis of pure possibilities.

APPENDIX: ITEMS OF PROPOSITIONAL REASONING TASK

1. Peter: Jim has the card with the star on it. Jim: Peter has the card with the star on it. Peter is lying; Jim is lying.

Correct answer: Impossible. (4 transformations)

2. Peter: Jim does not have the card with the star on it. Jim: Peter has the card with the star on it. Peter is telling the truth; Jim is lying.

Correct answer: Impossible. (4 transformations)

3. Peter: Jim does not have the card with the star on it. Jim: I have the card with the star on it. One child is telling the truth and the other is lying, but we do not know who is telling the truth and who is lying.

Correct answer: Impossible. (Formal operations)

4. Peter: I have the card with the star on it. Jim: Peter does not have the card with the star on it. Peter is lying; Jim is telling the truth.

Correct answer: Jim. (3 transformations)

5. Peter: I do not have the card with the star on it. Jim: I do not have the card with the star on it. Peter is telling the truth; Jim is telling the truth.

Correct answer: Impossible. (2 transformations)

6. Peter: I do not have the card with the star on it. Jim: Peter does not have the card with the star on it. Peter is lying; Jim is lying.

Correct answer: Peter. (5 transformations)

7. Peter: Jim does not have the card with the star on it. Jim: I do not have the card with the star on it. Peter is telling the truth; Jim is telling the truth.

Correct answer: Peter. (3 transformations)

8. Peter: I have the card with the star on it. Jim: I do not have the card with the star on it. Peter is lying; Jim is telling the truth.

Correct answer: Impossible. (2 transformations)

9. Peter: Jim has the card with the star on it. Jim: I have the card with the star on it. Peter is telling the truth; Jim is telling the truth.

Correct answer: Jim. (1 transformation)

10. Peter: Ann does not have the card with the star on it. Jim: Peter has the card with the star on it. Ann: I have the card with the star on it. Peter is telling the truth; Jim is lying; Ann is lying.

Correct answer: Jim. (5 transformations)

11. Peter: Ann has the card with the star on it. Jim: Ann does not have the card with the star on it. Ann: I do not have the card with the star on it. Peter is telling the truth; Jim is lying; Ann is lying.

Correct answer: Ann. (6 transformations)

12. Peter: Jim has the card with the star on it. Jim: Ann does not have the card with the star on it. Ann: Jim does not have the card with the star on it. Peter is lying; Jim is telling the truth; Ann is lying.

Correct answer: Impossible. (7 transformations)

13. Peter: I do not have the card with the star on it. Jim: Peter does not have the card with the star on it. Ann: Jim has the card with the star on it. Two children are telling the truth and one child is lying, but we do not know who is telling the truth and who is lying.

Correct answer: Ann. (Formal operations)

REFERENCES

Barouillet, P., & Lecas, J.-F. (1998). How can mental models theory account for content effects in conditional reasoning? A developmental perspective. *Cognition,* **67,** 209–253.

Beilin H., & Lust, B. (1975). A study of the development of logical and linguistic connectives. In H. Beilin (Ed.), *Studies in the cognitive basis of language development* (pp. 218–337). New York: Academic Press.

Bond, T. (1995a). Piaget and measurement I: The twain really do meet. *Archives de Psychologie,* **63,** 71–87.

Bond, T. (1995b). Piaget and measurement II: Empirical validation of the Piagetian model. *Archives de Psychologie,* **63,** 155–195.

Byrnes, J. P. (1988). Formal operations: A systematic reformulation. *Developmental Review,* **8,** 66–87.

Campbell, R. L. (1991). Does class inclusion have mathematical prerequisites? *Cognitive Development,* **6,** 169–194.

Campbell, R. L., & Richie, D. M. (1983). Problems in the theory of developmental sequences. *Human Development,* **26,** 156–172.

Carpendale, J. I., McBride, M. L., & Chapman, M. (1996). Language and operations in children's class inclusion reasoning: The operational semantic theory of reasoning. *Developmental Review,* **16,** 391–415.

Chapman, M. (1988). *Constructive evolution.* Cambridge: Cambridge Univ. Press.

Chapman, M., & McBride, M. L. (1992). Beyond competence and performance: Children's class inclusion strategies, superordinate class cues, and verbal justifications. *Developmental Psychology,* **28,** 319–327.

Davidson, P. M. (1987). Developing coordination of class unions, intersections, and complements. *Cognitive Development,* **2,** 21–36.

Dean, A. L., Chabaud, S., & Bridges, E. (1981). Classes, collections, and distinctive features: Alternative strategies for solving class inclusion problems. *Cognitive Psychology,* **13,** 84–112.

Elliott, C. (1982). The measurement characteristics of developmental tests. In S. Modgil & C. Modgil (Eds.), *Jean Piaget: Consensus and controversy* (pp. 241–255). London: Holt, Rinehart & Winston.

Englebretsen, G. (1981). *Logical negation.* Assen: Van Gorcum.

Flavell, J. H. (1963). *The developmental psychology of Jean Piaget.* New York: Van Nostrand.

Fox, C. M., & Gray, W. M. (1997, June). *An analysis of the ordering of different forms of thought and different logical operations.* Poster presented at the Twenty-Seventh Annual Symposium of the Jean Piaget Society, Santa Monica, CA.

Fuson, K. C., Lyons, B. G., Pergament, G. G., Hall, J. W., & Kwon, Y. (1988). Effects of collection terms on class inclusion and on number tasks. *Cognitive Psychology,* **20,** 96–120.

Girotto, V., Blaye, A., & Farioli, F. (1989). A reason to reason: Pragmatic basis of children's search for counterexamples. *European Bulletin of Cognitive Psychology,* **9,** 297–321.

Gold, R. (1986). Class inclusion performance: Effect of intermingling the subsets. *British Journal of Developmental Psychology,* **5,** 343–346.

Haars, V. J. E., & Mason, E. J. (1986). Children's understanding of class inclusion and their ability to reason with implication. *International Journal of Behavioral Development,* **9,** 45–63.

Hodkin, B. (1981). Language effects in the assessment of class-inclusion ability. *Child Development,* **52,** 470–478.

Hodkin, B. (1987). Performance model analysis in class inclusion: An illustration with two language conditions. *Developmental Psychology,* **23,** 683–689.

Inhelder, B., & Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence.* New York: Basic Books. (Original work published 1955)

Inhelder, B., & Piaget, J. (1969). *The early growth of logic in the child.* New York: Norton. (Original work published 1959)

Inhelder, B., Sinclair, H., & Bovet, M. (1980). *Learning and the development of cognition.* Cambridge, MA: Harvard Univ. Press. (Original work published 1974)

Jou, J. (1988). The comprehension of double negation in Chinese children. *Journal of Experimental Child Psychology,* **45,** 457–471.

Kant, I. (1974). *Logic.* New York: Dover. (Original work published 1800)

Kesselring, T. (1981). *Entwicklung und Widerspruch. Ein Vergleich zwischen Piaget's genetischer Erkenntnistheorie und Hegels Dialektik* [Development and contradiction. A comparison between Piaget's genetic epistemology and Hegel's dialectic]. Frankfurt am Main: Suhrkamp.

Kesselring, T. (1985). *Die Produktivität der Antinomie. Hegels Dialektik im Lichte der genetischen Erkenntnistheorie und der formalen Logik* [The productivity of antinomy. Hegel's dialectic in the light of genetic epistemology and formal logic]. Frankfurt am Main: Suhrkamp.

Kesselring, T. (1993). Egocentrism and equilibration. In D. Maurice & J. Montangero (Eds.), *Equilibrium and equilibration* (Cahiers de la Fondation Archives Jean Piaget, Vol. 12). Geneva: Fondation Archives Jean Piaget.

Kingma, J., & Van den Bos, K. P. (1989). Unidimensional scales: New methods to analyze the sequences in concept development. *Genetic, Social and General Psychology Monographs,* **114,** 477–508.

Markovits, H. (1993). The development of conditional reasoning: A Piagetian reformulation of mental models theory. *Merrill-Palmer Quarterly,* **39,** 131–158.

Markovits, H., Fleury, M.-L., Quinn, S., & Venet, M. (1998). The development of conditional reasoning and the structure of semantic memory. *Child Development,* **69,** 742–755.

Markovits, H., Venet, M., Janveau-Brennan, G., Malfait, N., Pion, N., & Vadeboncoeur, I. (1996). Reasoning in young children: Fantasy and information retrieval. *Child Development,* **67,** 2857–2872.

Matalon, B. (1990). A genetic study of implication. In W. F. Overton (Ed.), *Reasoning, necessity, and logic: Developmental perspectives* (pp. 87–110). Hillsdale, NJ: Erlbaum. (Original work published 1962)

McGarrigle, J., Grieve, R., & Hughes, M. (1978). Interpreting inclusion: A contribution to the study of the child's cognitive and linguistic development. *Journal of Experimental Child Psychology,* **26,** 528–550.

Morf, A. (1957). Les relations entre la logique et le langage lors du passage du raisonnement concret au raisonnement formel [The relationship between logic and language during the transition from concrete to formal reasoning]. In L. Apostel, B. Mandelbrot, & A. Morf (Eds.), *Logique, langage, et theorie de l'information* (pp. 173–204). Paris: Presses Universitaires de France.

Moshman, D. (1979). Development of formal hypothesis-testing ability. *Developmental Psychology,* **15,** 104–112.

Müller, U. (1999). Structure and content of formal operational thought: An interpretation in context. *Archives de Psychologie,* **67,** 21–35.

Müller, U., & Overton, W. F. (1998). How to grow a baby: A reevaluation of image-schema and Piagetian action approaches to representation. *Human Development,* **41,** 71–111.

Müller, U., Sokol, B., & Overton, W. F. (1998). Reframing a constructivist model of the development of mental representation: The role of higher-order operations. *Developmental Review,* **18,** 155–201.

Noelting, G., Coudé, G., Fiset, C., & Brunel, M.-L. (1993, June). *Logic as resulting from the dialectical interaction between operative subject and constraining field.* Paper presented at the 23rd Jean Piaget Society Symposium, Philadelphia, PA.

O'Brien, D. P. (1987). The development of conditional reasoning: An iffy proposition. *Advances in Child Development and Behavior,* **20,** 61–90.

O'Brien, D. P., & Overton, W. F. (1980). Conditional reasoning following contradictory evidence: A developmental analysis. *Journal of Experimental Child Psychology,* **30,** 44–61.

O'Brien, D. P., & Overton, W. F. (1982). Conditional reasoning and the competence–performance issue: A developmental analysis of a training task. *Journal of Experimental Child Psychology,* **34,** 274–290.

Overton, W. F. (1990). Competence and procedures: Constraints on the development of logical reasoning. In W. F. Overton (Ed.), *Reasoning, necessity, and logic: Developmental perspectives* (pp. 1–30). Hillsdale, NJ: Erlbaum.

Overton, W. F. (1998). Developmental psychology: Philosophy, concepts, and methodology. In R. M. Lerner (Ed.), *Theoretical models of human development: Vol. 1. Handbook of child psychology* (5th ed., pp. 107–188). New York: Wiley.

Piaget, J. (1950). *The psychology of intelligence.* New York: Harcourt, Brace. (Original work published 1947)

Piaget, J. (1952a). Autobiography. In E. G. Boring, H. S. Langfeld, H. Werner, & R. M. Yerkes (Eds.), *A history of psychology in autobiography* (Vol. 4, pp. 237–256). Worcester, MA: Clark Univ. Press.

Piaget, J. (1952b). *The child's conception of number.* London: Routledge & Kegan Paul. (Original work published 1941)

Piaget, J. (1963). *The origins of intelligence in children.* New York: Norton. (Original work published 1936)

Piaget, J. (1967). Genesis and structure in the psychology of intelligence. In J. Piaget, *Six psychological studies* (pp. 143–158). New York: Random House.

Piaget, J. (1971). *Structuralism.* London: Routledge & Kegan Paul. (Original work published 1968)

Piaget, J. (1972). *Essai de logique operatoire* [Essay on operatory logic]. Paris: Dunod.

Piaget, J. (1977a). *Recherches sur l'abstraction réfléchissante: 1. L'abstraction des relations logico-arithmétiques* [Experiments on reflective abstraction: Vol. 1. Abstraction of logico-arithmetic relations]. Paris: Presses Universitaires de France.

Piaget, J. (1977b). Some recent research and its link with a new theory of groupings and conservations based on commutability. *Annals of the New York Academy of Sciences,* **291,** 350–358.

Piaget, J. (1977c). The stages of intellectual development in childhood and adolescence. In H. E. Gruber & J. J. Vonéche (Eds.), *The essential Piaget* (pp. 814–819). New York: Basic Books. (Original work published 1955)

Piaget, J. (1978). *Success and understanding.* London: Routledge & Kegan Paul. (Original work published 1974)

Piaget, J. (1980a). *Experiments in contradiction.* Chicago: Univ. of Chicago Press. (Original work published 1974)

Piaget, J. (1980b). *Recherches sur les correspondances* [Experiments on correspondences]. Paris: Presses Universitaires de France.

Piaget, J. (1985). *The equilibration of cognitive structures.* Chicago: Univ. of Chicago Press. (Original work published 1975)

Piaget, J. (1987). *Possibility and necessity: Vol. 1. The role of possibility in cognitive development.* Minneapolis, MN: Univ. of Minnesota Press. (Original work published 1981)

Piaget, J. (1995). *Sociological studies.* London: Routledge. (Original work published 1965)

Piaget, J., & Garcia, R. (1991). *Toward a logic of meaning.* Hillsdale, NJ: Erlbaum.

Piéraut-Le Bonniec, G. (1980). *The development of modal reasoning: Genesis of necessity and possibility notions.* New York: Academic Press.

Reichenbach, H. (1947). *Elements of symbolic logic.* Toronto: Macmillan.

Richard, J. F., & Leynet, M. E. (1994). The inferential structure of class-inclusion tasks. *British Journal of Developmental Psychology,* **12,** 209–233.

Riesz, E. D., & Cantor, G. N. (1985). Piaget's grouping II: The model, its assessment, and the effect of training. *Developmental Review,* **5,** 289–308.

Schröder, E., Edelstein, W., & Hoppe-Graff, S. (1991). Qualitative analyses of individual difference in intraindividual change: Examples from cognitive development. In D. Magnusson, L. R. Bergman, G. Rudinger, & B. Törestad (Eds.), *Problems and methods in longitudinal research* (pp. 166–189). Cambridge: Cambridge Univ. Press.

Shipley, E. F. (1979). The class inclusion task: Question form and distributive comparisons. *Journal of Psycholinguistic Research,* **8,** 301–331.

Siegler, R. S. (1996). A grand theory of development. *Monographs of the Society for Research in Child Development,* **61** (1–2, Serial No. 246), 266–275.

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

Smith, L. (1982). Class inclusion and conclusions about Piaget's theory. *British Journal of Psychology,* **73,** 267–276.

Smith, L. (1991). Age, ability, and intellectual development in Piagetian theory. In M. Chandler & M. Chapman (Eds.), *Criteria for competence* (pp. 69–91). Hillsdale, NJ: Erlbaum.

Smith, L. (1992). Judgements and justifications: Criteria for the attribution of children's knowledge in Piagetian research. *British Journal of Developmental Psychology,* **10,** 1–23.

Smith, L. (1993). *Necessary knowledge. Piagetian perspectives on constructivism.* Hillsdale, NJ: Erlbaum.

Smith, L. (1999). Necessary knowledge in number conservation. *Developmental Science,* **2,** 23–27.

Thomas, H., & Horton, J. J. (1997). Competency criteria and the class inclusion task: Modeling judgment and justification. *Developmental Psychology,* **33,** 1060–1073.

Werner, H. (1957). The concept of development from a comparative and organismic point of view. In D. B. Harris (Ed.), *The concept of development: An issue in the study of human behavior* (pp. 125–148). Minneapolis: Univ. of Minnesota Press.

Winer, G. A. (1978). Enhancement of class-inclusion reasoning through verbal context. *Journal of Genetic Psychology,* **132,** 299–306.

Winer, G. A. (1980). Class-inclusion reasoning in children: A review of the empirical literature. *Child Development,* **51,** 309–328.

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

Wright, B. D., & Masters, G. N. (1982). *Rating scale analysis.* Chicago: MESA Press.

Wright, B. D., & Stone, M. H. (1979). *Best test design.* Chicago: MESA Press.

Received June 24, 1997; revised May 11, 1999