

Functions, Operations, and Decalage in the Development of Transitivity

Michael Chapman and Ulman Lindenberger

Max Planck Institute for Human Development and Education, Berlin, Federal Republic of Germany

The hypothesis that some attempts to reduce the performance demands of concrete operational tasks may have allowed children to solve those tasks with preoperational functions was tested by administering two previously used versions of the transitivity task for length and weight to 120 children 6 to 9 years of age. In the standard Piagetian version of the task, comparison objects were presented only two at a time, with no correlation between length or weight and spatial position. In the alternate version, objects were ordered in space by length or weight, but placed too far apart for differences in length to be seen. As predicted: (a) the standard version was solvable only through an operational composition of premise relations, but the alternate version allowed children to recode length or weight as a function of spatial position; (b) the functional solution was easier and developed at an earlier age than the operational solution for length as for weight. In addition, the decalage between length and weight was found to result from children's tendency to infer weight as a function of size rather than from an operational composition of premise relations.

One of the most problematic issues in Piagetian theory is the phenomenon of horizontal decalage—the fact that certain cognitive tasks presumably having the same logical structure are nevertheless solved by children at different ages. For many commentators, the existence of such horizontal decalages has cast doubt on the structural basis of Piagetian theory (e.g., see Bruner, 1983; Fischer, 1980).

Two types of developmental decalages have generally been discussed in this connection: decalages between tasks involving different contents or different procedures. *Content decalages* occur when the same logical or mathematical operations are found to develop at different ages with respect to different types of content. For example, Piaget and Inhelder (1941/1974) reported that conservation and transitivity developed at different ages with respect to physical quantity, weight, and volume. In contrast, *procedural decalages* result from certain modifications in the procedures by which Piagetian tasks are administered. Typically, tasks are modified in order to eliminate certain “performance factors” that are presumably irrelevant to the competence being assessed, and children are found to solve those tasks at an earlier age than that reported by Piaget (see review by Gelman & Baillargeon, 1983). Such results have frequently been interpreted as implying that Piaget's methods were too “conservative” and that he subsequently “underestimated” children's true abilities (see Brainerd, 1973a, 1977).

The present study was based on the premise that some pre-

viously used procedures designed to eliminate presumably extraneous performance factors in concrete-operational tasks may have actually altered those tasks in such a way as to allow children to solve them with functional, as opposed to operational, reasoning (Piaget, Grize, Szeminska, & Vinh Bang, 1968/1977). As defined by Piaget (1949/1972, 1957), concrete operational thinking involves a (reversible) composition of elements, which may include operations of classification and ordering. In contrast, the preoperational “constitutive functions” studied by Piaget et al. (1968/1977) involve only a one-way mapping of one logical variable onto another; no composition of discrete elements is implied.

Transitivity in Young Children

The foregoing hypothesis was investigated in the present study with respect to transitive reasoning. In Piagetian theory, transitive inference is defined in terms of the “addition of asymmetrical relations” and is a typical example of concrete operational thought (Piaget, 1949/1972, 1957). Like most other concrete operational tasks, transitive reasoning was found to appear in development only around the age of 7 to 8 years (Piaget, Inhelder, & Szeminska, 1948/1960). However, investigators using procedures designed to eliminate extraneous performance factors in Piaget's methods have reported transitivity in children 5 years of age and less. For example, Braine (1959) argued that Piaget's clinical interview method might require too much of children's verbal abilities. Using a nonverbal assessment procedure, he found children as young as 5 years of age to solve the transitivity task successfully. Similarly, Bryant and Trabasso (1971) argued that children might fail transitivity tasks not because they cannot reason, but because they forget the premises of a transitivity problem before they can draw the appropriate inference. In order to avoid such memory deficits, they trained children on the premises of transitive inferences and found that even 4-year-olds could solve the respective problems under these conditions. A number of subsequent researchers also re-

The research described in this article was conducted as part of a Diplomarbeit by Ulman Lindenberger in the context of the “Action, Control, and Task-performance” project at the Max Planck Institute for Human Development and Education (the principal investigators were Michael Chapman, Ellen A. Skinner, and Paul B. Baltes).

We would like to thank the following people for their valuable assistance: Peter Usinger for organization, Birgit Herbeck for school contacts, and Alexander von Eye for statistical consultation.

Correspondence regarding this article should be addressed to Michael Chapman, Department of Psychology, University of British Columbia, 2136 West Mall, Vancouver V6T 1Y7, Canada.

ported transitive reasoning in preschool children (see reviews by Breslow, 1981; Halford, 1982; Thayer & Collyer, 1978). Although Piaget denied that age was a criterion for his cognitive developmental stages (e.g., Piaget in Bringuier, 1980), these results were apparently discrepant with his own findings and cast doubt on both his explanations of transitive reasoning and his methods for assessing it.

Such reports of precocious transitivity in young children did not go unchallenged. Braine's methods were criticized by Smedslund (1963, 1965, 1966), who argued that children could solve this version of the task through a nontransitive labeling strategy (cf. Braine, 1964). This argument was apparently supported by studies designed to control for such strategies (Murray & Youniss, 1968; Youniss & Murray, 1970), but Brainerd (1973a) pointed out that the controls used in these studies might create problems of their own and that the importance of the nontransitive solutions described by Smedslund therefore remained to be demonstrated.

Early critics of Bryant and Trabasso's training paradigm argued that this version of the task might be solved by various nontransitive strategies, including labeling (De Broysson-Bardies & O'Regan, 1973) or "sublogical" inference (Youniss & Furth, 1973). According to Trabasso (1975, 1977), however, even these alternative strategies assume that children mentally represent the comparison objects in an ordered series. Thus, even if transitive inference is not required in the test phase of the experiment (because children can judge the relative lengths of any two objects simply by noting their positions in the series), transitivity would nevertheless be required in the training phase (because the ordered series constructed by the children has transitive properties). Against this argument, Breslow (1981) advanced a "sequential-contiguity" model to explain how an ordered series might be formed without transitive inference. But Halford and Kelly (1984) pointed out that an ordered series with transitive properties cannot be formed from relations of contiguity, because this relation is in fact nontransitive. Accepting Trabasso's argument that the true test of transitivity is the integration of premise relations during training, they presented data suggesting that most children under 5 years were incapable of such integration (see also Halford, 1984).

The suggestion that children might solve Bryant and Trabasso's version of the transitivity task through a labeling strategy was pursued in a series of studies by Perner and colleagues (Perner & Aebi, 1985; Perner & Mansbridge, 1983; Perner, Steiner, & Stachelin, 1981). Briefly, they found that young children's integration of premise comparisons into a linear series depended on specific visual or procedural cues regarding the gradation of length in the series (see also Kallio, 1982). Without such cues, children tended to label a given stick according to how often it had been "long" or "short" relative to other sticks in the premise comparisons. These results led Perner and Aebi (1985) to conclude that the skills used by children to solve Bryant and Trabasso's training task had little to do with transitive reasoning as such, a conclusion apparently supported by McGonigle and Chalmers' (1977) finding that squirrel monkeys were able to solve transitivity training tasks as well as young children (see also Chalmers & McGonigle, 1984). Instead of merely eliminating young children's memory deficits as a potential source of error in transitive reasoning, Bryant and Tra-

basso's training procedure may have significantly changed the nature of the task.

Other critics of Bryant and Trabasso investigated the role of memory in transitive reasoning by testing the statistical dependency between correct solutions and memory for premises. Thus, Halford and Galloway (1977) and Russell (1981) found that some children did not draw transitive inferences even though they recalled the necessary premises, suggesting that failure cannot be explained solely through memory deficits. However, Brainerd and Kingma (1984) pointed out that a true test of the proposed dependence of transitive reasoning on memory for premises involves comparing (a) the unconditional probability of a correct inference with (b) the conditional probability of a correct inference given memory for premises (and, inversely, comparing the unconditional probability of recalling the premises with the conditional probability of recalling the premises given a correct inference). In a reanalysis of Halford and Galloway's and Russell's data as well as in new experiments of their own, Brainerd and Kingma found that transitive inference and memory for premises were in fact stochastically independent of each other. They argued that these results contradicted Piagetian as well as most information-processing models of transitivity performance, which assume that memory for premises is necessary for transitive inference. As an alternative, they proposed a "fuzzy trace" theory of children's reasoning and memory, according to which children solve transitivity tasks through imprecise memory traces of the form, for example, "Things get bigger to the right."

In fact, the latter formula is a good example of what Piaget et al. (1968/1977) called a "constitutive function"; relations of length are inferred as a function of spatial relations. According to this interpretation, the children in Brainerd and Kingma's study solved the "transitivity" problem through functional reasoning of this kind rather than from an operational composition of premise relations. Correct solutions were stochastically independent of memory for premises, because correct answers could be inferred without considering those premises at all. (On the prevalence of relational thinking in young children, see also Bryant, 1974.)

The analysis of this task in terms of functional reasoning, however, suggests some limitations on this type of solution: Children would be able to infer relations of length as a function of spatial relations only under the condition that the comparison objects were in fact ordered in space according to length. Thus, functional reasoning cannot be considered a *general* model of performance on transitivity tasks, but it could provide a viable model of performance in many studies in which comparison objects were ordered in space. In those studies at least, it may account for the occurrence of precocious "transitivity" among preschoolers.

Standard Versus Alternate Transitivity Tasks

The foregoing analysis was investigated in the present study through a comparison between two versions of the transitivity task, one designed to be solvable only through operational composition of premise relations and the other to be solvable (at least potentially) by functional reasoning. By convention these two tasks will be called the *standard* and *alternate* versions, re-

spectively. In neither case were children trained on premise relations, because the explicit object of the present investigation was transitive reasoning as it occurs without such training.

The standard version was based on procedures described by Piaget (1967, p. 50) and was designed to ensure that a correct solution could be inferred only from a composition of individual premise relations. Thus, both premise comparisons and transitivity problems were presented by showing the comparison objects only two at a time, with the remaining objects hidden out of sight. Care was taken to avoid any correlation between length and spatial (right-left) position so that the objective conditions theoretically necessary for functional reasoning would be avoided.

The alternate version was based on the procedure used by Brainerd (1973b) and Brainerd and Kingma (1984, 1985), which in turn was a generalization of assessment methods used in many previous studies of transitivity (e.g., Braine, 1959; Brainerd, 1973c, 1979; Smedslund, 1963, 1966). Thus, comparison objects were arranged from left to right in order of increasing (or decreasing) length, but spaced far enough apart so that differences in length could not be seen directly. "Premise" comparisons were made by moving adjacent objects next to each other so that the difference in length became apparent. "Transitivity" questions referred to the relative lengths of the nonadjacent objects, which had not been directly compared.

Note that a similar correlation between length and spatial position existed in the test phase of Bryant and Trabasso's (1971) study as well as in many (but not all) of the subsequent investigations based on that study. As Trabasso (1975, p. 160) wrote of the original procedure, "a display box was in full view and the colored sticks were presented so that their locations in the box were correlated with their spatial location in the full linear order." Even though sticks were only compared two at a time outside the box, the serial order of sticks in the box was apparent (cf. the representation of Bryant and Trabasso's display box in Cohen, 1983, p. 113).

The assumption of investigators using methods analogous to the alternate version has generally been that children judge nonadjacent comparisons (e.g., A vs. C) through some integration of the premise comparisons (e.g., A vs. B and B vs. C). But according to the present analysis, an objective correlation between length and spatial position should make possible a functional solution (the fact that A is longer than C being inferred from the fact that A is to the left of C). In particular, the following predictions were made: (a) children should prefer to give functional justifications on the alternate version even if they were capable of operational justifications on the standard version; (b) children should be able to solve the alternate version at a higher rate and at an earlier age than the standard version, because functional reasoning is less difficult and developmentally prior to operational reasoning; (c) variation in the number of comparison objects should affect performance only in the standard version, because only the standard version requires that each of the pairwise premise comparisons actually be considered; in the alternate version, the transitivity question should be answerable simply by noting which of the two objects to be compared is further to the right, no matter how many intermediate objects intervene.

Length Versus Weight

A secondary purpose of this study was to examine the role of functional reasoning in the content decalage between length and weight. According to Piaget and Inhelder (1941/1974), this decalage results from the fact that children capable of operational transitivity with respect to length still tend to infer relations in weight from relations in size or undifferentiated "bigness" (see also Piaget, 1974). Thus, some children who are capable of operational composition in the transitivity-of-length task will nevertheless tend to derive relative weight as a function of relative size in the transitivity-of-weight task under the condition that the size and weight of comparison objects vary independently. For both versions of the transitivity task it was predicted (a) that children would demonstrate transitivity of length at an earlier age on the whole than transitivity of weight and (b) that children would tend to fail the transitivity-of-weight task because of a tendency to infer relations of weight as a function of relations in size.

Study 1

Method

Subjects and procedures. A total of 120 first, second, and third graders, equally divided by sex and grade, were investigated in two elementary schools from a middle- and lower-middle-class area of West Berlin. Mean ages of first, second, and third graders were 7.06, 7.91, and 8.95 years, respectively. In a single 45-min session, all children were administered both standard and alternate versions of the transitive-reasoning task for length and weight problems. The number of comparison objects in both versions was varied from three to five (except in the standard version for weight, in which only three and four objects were used). Tasks were administered in the following order: three-term standard length task, three-term standard weight task, five-term standard length task, four-term standard weight task, four-term standard length task, alternate length task (three to five terms), and alternate weight task (three to five terms). This order was chosen so that possible carry-over effects would work against the major hypotheses of the study: Carry-over from the standard tasks to the alternate tasks would, if anything, maximize the probability that children would use the same form of justification on the alternate as on the standard tasks. In presenting premise comparisons and in asking about nonadjacent relations, only single comparatives were used ("longer" or "heavier"). Details of administration were as follows:

Standard length tasks. A black stick A (length = 14.1 cm), a light-brown stick B (14.4 cm), and a dark-brown stick C (14.7 cm) were used. At the beginning of the problem all three sticks were displayed with the ends hidden so that their relative lengths could not be perceived. The children were asked to name the color of the sticks. Then the experimenter removed all sticks and placed them in a box under the table. Sticks B and C were reintroduced and placed upright on the table next to each other, and children were asked to indicate which stick was longer. Then both sticks were removed and the same was done with sticks A and B. Memory for these "premise comparisons" was checked by repeating the preparatory questions twice, this time with the ends of the sticks hidden to conceal their differences in length. Children's answers were corrected (through visual comparison) the first time. Then sticks A and C were displayed with the ends hidden so that no difference in length was visible, and children were asked which stick was longer. Finally, children were asked to explain their judgments.

In the standard length task with four comparison objects, four sticks of different colors varying in length from 13.8 cm to 14.7 cm (with .3

cm difference) were used. The order of preparatory questions was AB, CD, BC. In the five-stick version, sticks varied from 13.8 cm to 15.0 cm, and the order of preparatory questions was BC, DE, CD, AB. In other respects, these tasks were administered in the same way as the three-object version.

Standard weight tasks. All weight tasks were administered by means of a balance scale manipulated by the experimenter; children were not allowed to touch the scale nor the balls being weighed. To familiarize children with the use of a balance, the experimenter introduced two brown balls of identical diameters but different weights (200 g and 260 g, respectively), placed them on the balance, and asked the children to indicate the heavier ball. When it was clear that children understood the principle that the heavier ball always went down, both balls were removed from the table. Three other balls were introduced and shown to the children, a silver ball A (weight = 25 g, diameter = 7.5 cm), a black ball B (180 g, 5.5 cm) and a gray ball C (340 g, 6.5 cm). Preparatory questions referred to balls AB and BC, respectively. As a memory check, these questions were repeated twice without weighing the balls, and children's mistakes were corrected (through visual comparison with the scale) the first time. Then the experimenter held balls A and C in his hands and asked the children which ball was heavier. Finally, children were asked to explain their judgments.

The standard weight task with four comparison objects featured four balls, a red ball A (35 g, 6.5 cm), a green ball B (100 g, 9.0 cm), a white ball C (180 g, 5.5 cm), and a yellow ball D (340 g, 6.5 cm). The order in which the preparatory questions was asked was BC, CD, AB. In other respects, the task was administered in the same way as the three-object task.

Alternate length task. Five sticks A, B, C, D, and E were used, varying in length from 13.8 cm to 15.0 cm, with a difference of 0.3 cm between each successive stick. All sticks were of the same color (purple). At the beginning of administration, the sticks were placed in ascending order on the table at a distance of roughly 30 cm between adjacent sticks so that it was impossible to perceive the respective lengths of the sticks directly. Preparatory questions referred to the relative lengths of adjacent sticks in the order AB, BC, CD, and DE. In presenting these "premise" comparisons, the experimenter placed each pair of adjacent sticks next to each other so that the length difference became visible and asked children to indicate the longer stick. Then the experimenter returned the sticks to their original positions and tested memory for premise comparisons by asking about adjacent sticks twice in the same order as before. Wrong answers were corrected (by moving the sticks together again) the first time. Test questions referred to the relative lengths of nonadjacent sticks, which had not been seen together. Sticks remained spaced apart from each other during this test phase. Questions about sticks AC, BD, and CE are referred to as three-term problems, AD and BE as four-term problems, and AE as a five-term problem, referring to the number of sticks involved in each case. After children answered the test questions, they were asked to explain their judgments.

Alternate weight task. Five balls of the same color (blue) were used: ball A (35 g, 7.5 cm), ball B (90 g, 6.0 cm), ball C (100 g, 9.0 cm), ball D (200 g, 7.5 cm), and ball E (250 g, 6.5 cm). At the beginning of the problem, the balls were placed on the table in ascending order, with the lightest ball to the left and the heaviest ball to the right. The distance between adjacent balls was roughly 30 cm. The "premise comparisons" referred to the relative weights of adjacent balls. The order of these comparisons was AB, BC, CD, and DE. To present the weight relations between adjacent balls, the experimenter placed each pair of adjacent balls on the balance scale and asked children to indicate the heavier ball. Memory for premise comparisons was tested by repeating the questions about adjacent balls twice in the same order without weighing the balls. Wrong answers were corrected (using the scale) the first time. Test questions referred to the relative weight of nonadjacent balls and were posed in the following order: BD, AC, BE, AD, CE, and AE. The balls were

not weighed in the scale during this test phase. AC, BD, and CE are referred to as three-term problems, AD and BE as four-term problems, and AE as a five-term problem. After the children had answered the test questions they were asked to explain their judgments.

Coding

We considered children to have passed a task when a correct judgment was followed by a valid explanation, and we accepted both verbal and nonverbal judgments (i.e., pointing). The following two types of explanations were considered: (a) In operational explanations, judgments were justified in terms of a comparison of premise relations. For example, "this one (C) is longer, because it's longer than that one (B), and that one (B) is longer than the other one (A)." (b) In functional explanations, judgments about relations in length or weight were justified in terms of spatial position or some other single dimension (e.g., size). For example, "they get bigger in this direction," or "this one (C) is bigger than that one (A), because it's on the right." On the alternate tasks, both types of explanations could be valid, because procedures allowed a correct judgment to be derived either as a function of spatial position or from the operational comparison of relations. On the standard tasks, only an operational explanation was valid, because the procedure ruled out any correlation between relations in length or weight and spatial position. In either case, operational explanations were considered valid only if all necessary premise relations were mentioned. Interrater agreement in coding operational and functional explanations was 95.0% and 95.8%, respectively. In order to investigate the reasons for the length-weight decalage, children's erroneous explanations in the standard weight tasks were classified into categories. Interrater reliability for these categories was 90%.

Recall that children were required to give multiple judgments in three- and four-term alternate tasks (e.g., AC, BD, and CE judgments in the three-term task). In these tasks, they were credited with a "pass" only if they gave a correct judgment on all comparisons, followed by a valid explanation. Occasionally (in less than 5% of all answers), children gave both functional and operational explanations on the alternate tasks. In such cases, only the preferred (i.e., the first) explanation was counted because the question of interest was one of performance, not competence (i.e., which form of reasoning children preferred to use even if they were capable of both).

Results

Alternate versus standard tasks. The percentage of children in each grade passing the various tasks is presented in Table 1. As indicated in the note to this table, all children passing the alternate version of the transitivity tasks gave functional justifications as their preferred explanations. The binomial probability of such an outcome under the a priori assumption of equal probability for the two types of explanations was 1.50×10^{-36} for length and 7.11×10^{-15} for weight.

Other predictions regarding the differences between standard and alternate procedures were also supported by the data shown in Table 1. The number of comparison objects did not affect performance in the alternate version of the transitivity tasks. All children solving three-term problems solved four- and five-term problems as well. In the standard version, however, the number of comparison objects greatly affected performance. For length tasks, differences in percent correct between three-, four-, and five-term problems were statistically significant for first graders, Cochran's $Q(2) = 18.20$, for second graders, $Q(2) = 33.60$, and for third graders, $Q(2) = 35.27$ (all $ps < .001$). For weight tasks,

Table 1
Study 1: Percent Correct for Alternate
and Standard Transitivity Tasks

Version of task	Grade		
	First	Second	Third
Transitivity of length			
Alternate ^a			
Three-, Four-, and Five-term	97.5	100.0	100.0
Standard			
Three-term	25.0	52.5	60.0
Four-term	2.5	12.5	20.0
Five-term	0.0	2.5	5.0
Transitivity of weight			
Alternate ^a			
Three-, Four-, and Five-term	20.0	47.5	50.0
Standard			
Three-term	7.5	35.0	42.5
Four-term	2.5	10.0	10.0

Note. $n = 40$ in each grade.

^a All children passing the alternate tasks gave functional explanations as their preferred explanations. All children solving three-term problems solved four- and five-term problems as well.

differences between three- and four-term problems were significant for second graders, $Q(1) = 10.00, p < .005$, and for third graders, $Q(1) = 13.00, p < .001$, but not for first graders, $Q(1) = 2.00$. This last result was attributable to the fact that only 3 first graders passed either of the tasks in question.

With respect to the procedural decalage between alternate and standard tasks, the alternate version tended to be solved by a higher percentage of children in all grades than the standard version, as was predicted. For length tasks, differences in percent correct between the alternate version and the standard three-term task were statistically significant for first graders, $Q(1) = 29.00$, for second graders, $Q(1) = 19.00$, and for third graders, $Q(1) = 16.00$ (all $ps < .001$). In fact, the alternate version and the standard three-, four-, and five-term tasks taken together formed a perfect Guttman scale in the total sample. Using the marginal distribution of passes and failures on each task to compute an a priori probability of scalable versus non-scalable profiles, the binomial probability of such an outcome was $p < 1.39 \times 10^{-5}$. For weight tasks, the differences between alternate and standard three-term tasks were significant for first graders, $Q(1) = 29.00, p < .001$, for second graders, $Q(1) = 5.00, p < .05$, but not for third graders, $Q(1) = 1.00$. Among the third graders, only 3 children solved the alternate task without also solving the three-term standard problem, and 1 child solved the latter task without solving the alternate task. Except for this 1 child, the alternate version together with the three- and four-term standard tasks also formed a Guttman scale in the total sample (cumulative binomial $p < 1.56 \times 10^{-12}$, given the marginal distribution of passes and failures).

Length versus weight. Turning to comparisons between the two content areas, the decalage between length and weight is clearly visible in Table 1. Focusing on the three-term standard

task, the difference in percent correct for length and weight was significant for first graders, $Q(1) = 5.44, p < .02$, for second graders, $Q(1) = 7.00, p < .01$, and for third graders, $Q(1) = 5.44, p < .02$. As indicated in Table 2, children's erroneous explanations on standard weight tasks were all functional in form. Relations of weight were inferred (a) as a direct function of size (e.g., "This one is heavier because it's bigger"), or (b) as an inverse function of size (e.g., "The smaller ones are heavier"), or (c) as a function of texture (e.g., "This one is heavier, because it looks harder"). Inferences of weight as a direct function of size tended to decrease across age. This trend was statistically significant for the three-term standard weight task, $\chi^2(2, N = 120) = 9.60, p < .01$, but only marginally significant for the four-term standard task, $\chi^2(2, N = 120) = 5.16, p < .10$. Because none of the other categories of errors decreased appreciably with age, the overall development in the operational transitivity of weight from first to third grade can be primarily attributed to this decrease in the tendency to infer weight as a direct function of size.

Discussion

In general, the results were consistent with the predictions. With respect to the procedural decalage between the alternate and standard transitivity tasks, children were found to solve the alternate version with functional reasoning, and the number of comparison objects affected performance only on the standard version. These results support the conclusion that the two tasks involved different logical structures. The spatial ordering of comparison objects in the alternate task allowed children to derive a correct answer using preoperational reasoning (length or weight derived as a function of spatial position), but the standard version required the operational composition of relations for its solution. Therefore, the developmental lag between the

Table 2
Study 1: Distribution of Errors in Standard
Transitivity of Weight Tasks

Type of explanation	Grade		
	First	Second	Third
Three-term task			
Weight a direct function of size	16	10	4
Weight an inverse function of size	0	4	2
Weight a function of texture	6	4	7
No explanation	12	8	10
Missing data	3	0	0
Total errors	37	26	23
Four-term task			
Weight a direct function of size	11	5	4
Weight an inverse function of size	0	3	0
Weight a function of texture	6	9	8
No explanation	20	19	24
Missing data	2	0	0
Total errors	39	36	36

Note. $n = 40$ in each grade.

two tasks must be considered an example of a structural stage difference rather than a true horizontal decalage.

In contrast, the content decalage between length and weight also found in this study can be considered an example of a genuine horizontal decalage, because in this case the same logical structure (operational composition of relations) appeared with high frequency only at two different points in development. Analysis of children's explanations suggested that the developmental lag between length and weight occurred because children tend to infer relative weight as a function of size even after they become capable of operational reasoning in regard to length. In short, perceived size acts as a misleading cue in regard to inferences involving weight but not in regard to inferences involving length.

In their experience with physical objects, children learn that bigger things are generally heavier. But they do not understand at first that this relation holds true only under special conditions—when the densities of the respective objects are equal. Thus, the understanding of weight as a content involves a coordination between size and density. Length as a content involves no such complexities. In Piaget's words, weight is a cognitively more complex content than a physical quantity such as length (Piaget, 1974; Piaget & Inhelder, 1941/1974).

If the content decalage between length and weight results from the misleading cue of physical size, however, the decalage between the transitivity of length and the transitivity of size would appear only when the comparison objects used in the weight task actually differed in size. This hypothesis was tested in Study 2.

Study 2

The purpose of Study 2 was twofold: first, to replicate the findings of Study 1 with respect to the procedural decalage between the standard and alternate versions of the transitivity task and, second, to test the hypothesis that the content decalage between the transitivity of length and the transitivity of weight would not appear when the comparison objects in the weight task do not differ in size.

Method

As part of a larger longitudinal study, we tested 27 of the 40 first graders and 32 of the 40 second graders from Study 1 a second time 1 year later when they were second and third graders, respectively. In addition, we tested 40 first graders who had not been tested before. Mean ages of first, second, and third graders were 7.07, 8.16, 8.89, respectively. As in Study 1, all children were administered standard and alternate versions of the transitive reasoning task for length and weight problems in a single sitting. Procedures followed those of Study 1, with one exception: The balls used as comparison objects in the transitivity-of-weight task were all of the same size (diameter = 6.5 cm). Their weights were as follows: for the alternate task, 50 g, 90 g, 170 g, 240 g, and 275 g; for the three-term standard task, 60 g, 170 g, and 270 g; for the four-term standard task, 55 g, 170 g, 220 g, and 395 g.

Results

Alternate versus standard tasks. The percentage of children in each grade passing the various tasks is presented in Table 3. Once again, all children passing the alternate version of the

Table 3
Study 2: Percent Correct for Alternate and Standard Transitivity Tasks

Version of task	Grade		
	First (<i>n</i> = 40)	Second (<i>n</i> = 27)	Third (<i>n</i> = 32)
Transitivity of length			
Alternate ^a			
Three-, Four-, and Five-term	100.0	96.3	100.0
Standard			
Three-term	37.5	59.3	65.6
Four-term	0.0	7.4	21.9
Five-term	0.0	0.0	6.3
Transitivity of weight			
Alternate ^a			
Three-, Four-, and Five-term	90.0	85.2	90.6
Standard			
Three-term	30.0	44.4	68.8
Four-term	5.0	14.8	25.0

^a All children passing the alternate tasks gave functional explanations as their preferred explanations. All children solving three-term problems solved four- and five-term problems as well.

transitivity tasks gave functional rather than operational explanations as the primary justification for their correct answers (binominal $p < 10^{-26}$, for both length and weight, under the assumption of equal a priori probability). The findings from Study 1 were replicated in other respects as well. The number of comparison objects affected performance only in the standard version. For length tasks, differences in percent correct between three-, four-, and five-term problems were statistically significant for first graders, $Q(2) = 30.00$, for second graders, $Q(2) = 28.50$, and for third graders, $Q(2) = 30.63$ (all $ps < .001$). For weight tasks, these differences were significant for first graders, $Q(1) = 10.00$, $p < .005$, for second graders, $Q(1) = 8.00$, $p < .005$, and for third graders, $Q(1) = 12.25$, $p < .001$.

The decalage between alternate and standard tasks is also clearly visible in Table 3. In all grades, the alternate version was solved by a higher percentage of children than the standard version. For length tasks, differences in percent correct between the alternate version and the standard three-term task were statistically significant for first graders, $Q(1) = 25.00$, $p < .001$, for second graders, $Q(1) = 10.00$, $p < .005$, and for third graders, $Q(1) = 11.00$, $p < .001$. For weight tasks, these differences were significant for first graders, $Q(1) = 24.00$, $p < .001$, for second graders, $Q(1) = 11.00$, $p < .001$, and for third graders, $Q(1) = 7.00$, $p < .01$. Both length and weight tasks formed perfect Guttman scales in the total sample. For length tasks, the binominal probability that the profiles of all 99 children in Study 2 would have conformed to a Guttman scale was .001, given the marginal distributions of successes and failures. For weight tasks, this probability was 1.4×10^{-6} .

Length versus weight. As shown in Table 3, differences in percent correct for the three-term standard length versus weight tasks were small, and none were statistically significant: $Q(1) =$

Table 4
 Study 2: Distribution of Errors in Standard
 Transitivity of Weight Tasks

Type of explanation	Grade		
	First (<i>n</i> = 40)	Second (<i>n</i> = 27)	Third (<i>n</i> = 32)
Three-term task			
Weight a direct function of size	10	5	3
Weight an inverse function of size	0	1	0
Weight a function of texture	6	2	1
Turn-taking	1	0	0
No explanation	11	7	6
Total errors	28	15	10
Four-term task			
Weight a direct function of size	10	6	5
Weight an inverse function of size	0	2	0
Weight a function of texture	7	2	0
Turn-taking	2	1	1
Three-term explanation	2	1	2
No explanation	17	11	16
Total errors	38	23	24

1.80, 2.67, and .20 for first, second, and third graders, respectively, and $Q(1) = 1.56$, for all three grades combined.

As shown in Table 4, nearly all of children's erroneous explanations on the standard transitivity-of-weight tasks could be classified into the same categories found in Study 1, namely, inferring relations in weight as a (direct or inverse) function of relations in size or as a function of texture. In addition, two other categories of errors were observed with low frequency: (a) Some children judged that a given ball was heavier "because it has never been heavier before." In effect, they argued that it was this particular ball's "turn" to be heavier. This category of explanation is labeled *turn-taking* in Table 4. (b) Some children justified their answers on the four-term task with a three-term operational explanation. That is, they argued that ball A was heavier than ball D, because A was heavier than B, and B was heavier than D, neglecting or forgetting to mention term C. Although this type of answer is operational in form, it is technically invalid because not all the premises necessary to infer the conclusion were mentioned.

In Table 4 one can also see that the decline in total errors was primarily attributable to declines in functional thinking (weight inferred as a direct function in size or as a function of texture). The decline in these two categories of functional thinking taken together from first to third grade was statistically significant for both the three-term and the four-term tasks. $\chi^2(2, N = 99) = 6.81$ and $\chi^2(2, N = 99) = 6.09$ (both $ps < .05$), respectively.

Discussion

With respect to the procedural decalage between the alternate and standard versions of the transitivity tasks, the following re-

sults of Study 1 were replicated: (a) Children solved the alternate version with greater frequency and at an earlier age than the standard task, (b) they solved the alternate task using functional reasoning, and (c) the number of comparison objects affected performance only on the standard version.

As for the comparison between length and weight, the magnitude of the developmental decalage between them was considerably reduced in Study 2, in which the balls were of equal size as compared with Study 1 (in which they differed in size). In fact, differences between performances on the length and weight tasks in Study 2 were not statistically significant. The fact that no significant length-weight decalage was found among first graders in Study 2 (who had not been tested a year before) rules out the effects of such repeated testing as an explanation for this finding.

Children who failed the weight tasks, however, did so primarily for the same reasons as in Study 1: Relations in weight tended to be inferred as a function of size or texture. Paradoxically, differences in size were sometimes given as a reason for differences in weight even when the comparison objects did not in fact differ in size. The most likely explanation for this finding is that children believed size differences were simply too small to be observed (just as length differences were too small to be observed when sticks were spaced out in the alternative length task). Alternatively, children might use the word *bigger* in an undifferentiated fashion to mean *more massive* and not simply to mean bigger in size (cf. Piaget, 1974, on the differentiation of "weight-as-heaviness" and "weight-as-mass").

General Discussion

The results of Studies 1 and 2 demonstrate that the procedural decalage between the standard and alternate versions of the transitivity task occurs because children can solve these tasks using two different inferential schemes: The standard version can be solved only through an *operational* composition of premise relations, but the alternate version can be solved through a *functional* derivation of relations in length or weight from relations in space. Children can solve the alternate version at an earlier age, because the use of such functional inferences develops before the use of operational compositions (Piaget et al., 1968/1977). Thus, the developmental lag between the two tasks must be considered a structurally defined stage transition (different structures emerging at different points in development) rather than a horizontal decalage as such (the same structure emerging at different points in development for different versions of the same task).

In fact, even children who were capable of operational reasoning on the standard version preferred to use functional reasoning on the alternate task. One explanation for this finding is that children confronted with optional methods of solving a problem may tend to choose the one that is cognitively less difficult. Given the option of solving a task with functional or operational reasoning, children tended to use the functional solution because it required less attentional capacity (Chapman, 1987a; on the role of capacity in children's reasoning, see Case, 1985; Halford, 1982; Pascual-Leone, 1970). Whereas the functional solution requires the coordination of only two relations (the spatial relation between comparison objects and the in-

ferred relation of length or weight), the operational solution requires a minimum of three relations (two "premise" relations and the "conclusion"). Thus, the functional and operational solutions differ both in structure (the form of the respective coordinations) and in attentional demand (the number of relations or representational schemes coordinated).

The capacity hypothesis also explains why the number of comparison objects affected performance only in the standard version and not in the alternate version. Although the three-term standard transitivity task involves a coordination of only three relations (the AB, BC, and AC comparisons), the four-term task requires four (AB, BC, CD, and AD). Thus, some children who are capable of solving the three-term task will fail the four-term task, not because they are incapable of "transitivity" as such, but only because they lack sufficient attentional capacity for solving the four-term task. In contrast, the functional solution of the alternate version requires only the coordination of two relations (e.g., the spatial relation between A and D and the relation of length or weight between A and D), no matter how many other terms intervene.

Functional reasoning was also implicated in the content decalage between length and weight. This decalage was found to occur primarily because children persisted in attempting to infer relations of weight as a function of relations in size or surface texture, even after they had become capable of operational compositions with regard to weight. These findings help to explicate what Piaget (e.g., 1971) called the "resistance" of certain contents to concrete operations: Weight "resists" operational compositions, because it can be inferred as a function of size with less mental effort. However, when children realize that the size-weight function leads to contradictions (i.e., bigger things are not always heavier), this function becomes "discredited" (Chapman, 1987a) and no longer provides a credible alternative to operational reasoning.

Transitivity in Young Children?

According to the present analysis, transitivity tasks that allow children to infer relative length as a function of spatial relations can be solved at a much earlier age than tasks requiring operational composition of premise relations. This interpretation raises the question of the extent to which the procedures used in those studies might have allowed children to solve transitivity tasks as a function of spatial cues.

In Braine's (1959) influential monograph on nonverbal methods in Piagetian research, for example, transitivity was assessed in the following manner: Sticks A and C were placed vertically on a table in front of the child and spaced far enough apart so that their difference in length was not apparent. They could be compared only through the use of a movable intermediary B. Braine assumed that children inferred the relative length of A and C from the AB and BC comparisons according to the standard formula for transitive reasoning. However, the fact that the spatial series ABC either ascended or descended in length from left to right meant that children would have been able to judge the relative lengths of A and C through position cues alone. From either the AB or the BC comparison children would have known whether the length series ascended or descended from right to left, and from this information they could

have inferred the relative length of A and C (e.g., $A < C$, because C is on the right, and things get longer in that direction). Consideration of both premise comparisons (i.e., operational composition) would not have been necessary. Ironically, Braine's use of nonverbal methods in an attempt to improve on the validity of Piaget's assessment procedures may have had the opposite effect, allowing children to solve a task meant to measure transitive inference by means of spatial position cues. Assessment procedures in which length was correlated with position were also used in other studies in which preschoolers were reported capable of transitive reasoning (e.g., Brainerd, 1973b, 1973c, 1979; Brainerd & Kingma, 1984, 1985).

In Piaget's original study of transitivity in spontaneous measurement (Piaget et al., 1948/1960), the object used as an intermediary in comparing two spatially separated towers was not necessarily intermediate in length. Thus, children were not confronted from the beginning with a spatial series in which length was correlated with spatial position, and they accordingly had no objective basis for functional reasoning on the basis of position cues. Smedslund (1963) used an assessment procedure comparable with Braine's, but he credited children with transitive reasoning only if they gave explanations that explicitly referred to both premise relations and might thereby have eliminated solutions solely on the basis of position cues. Murray and Youniss' (1968) procedure of assessing transitive inferences in a series involving both equalities and inequalities (e.g., $A > B = C$) would have eliminated the correlation between length and position, removing the objective conditions for functional reasoning on the basis of position cues (see also Youniss & Murray, 1970). Significantly, transitivity was not found among young children in these studies.

Although length was correlated with spatial position in the display box used by Bryant and Trabasso (1971), their results were replicated in subsequent studies in which such spatial cues were apparently eliminated (Kallio, 1982; Riley & Trabasso, 1974; Trabasso, 1975). Therefore, reports of transitivity in young children in studies following the Bryant and Trabasso training paradigm cannot be explained simply by the presence of such a correlation. The question is rather how children solve this type of task and whether it can be called "transitivity" in any meaningful sense.

On the basis of serial position and inferential distance effects obtained in the premise-training paradigm, Trabasso (1975, 1977) argued that children solve transitivity problems by recoding comparisons into a linear series with spatial properties so that any two members of the series can be compared directly. Their understanding of transitivity is reflected in their integration of premise relations into a single series with transitive properties. In several studies, however, such integration was found to occur only under specific task conditions: It did not occur if children were not provided with salient serial-order cues (Kallio, 1982) or other information regarding the gradation of length (Perner & Aebi, 1985). Under these conditions, premise comparisons were apparently represented separately. Independent representation of premise relations was also indicated in this study by the fact that, in contrast to the evidence cited by Trabasso (1975, 1977), the difficulty of the standard transitivity tasks increased rather than decreased with inferential distance (i.e., with the number of terms in a problem). Trabasso's serial-

representation model, therefore, cannot be considered a general model of transitive inference, but one pertaining only to specific conditions occurring in the premise-training paradigm. Instead of merely removing an extraneous source of measurement error, the procedure of training children on memory for premises may have changed the nature of the task.

Moreover, the present model of functional reasoning may provide an alternative model of children's performance in such experiments. Riley's (1976) replication of serial-order effects for nonphysical relations such as "happier than" or "nicer than" suggests that the serial representation formed by children in the premise-training paradigm need not in itself contain any information about a physical dimension such as length. Instead, children in the phase of premise training could (a) represent the correlation between length and serial position in the form of a length-position function and (b) use this function to translate relative length into relative position in a purely spatial representation. In the test phase, they could then retranslate relative position into relative length. In short, children would solve transitivity problems following training in much the same way that they solved the alternate tasks in the present study, except that the spatial series of comparison objects would have been interiorized through the training procedure. Such a form of representation would be economical, involving seriation along only a single (spatial) dimension, rather than two dimensions (length and spatial relations). The formation of such a purely spatial series would not imply any specific understanding of the transitivity of length as such, no more than children's use of spatial cues in the present alternate tasks implied any understanding of the transitivity of length.

Along with previous critics of the Bryant and Trabasso training procedure, we conclude that this paradigm has little to do with children's understanding of transitivity in the usual sense. Instead of merely removing an extraneous source of error variance, the training procedure enabled children to solve transitivity problems in a qualitatively different way than without training ("integration" of premise relations in a single serial representation through extended practice vs. "composition" of independently represented premise relations through a cognitive operation). The form of reasoning used by children in the learning paradigm may be of interest in its own right, but it should not be confused with different forms of reasoning used under different task conditions. Such confusion has occurred in part because researchers have tended to focus exclusively on children's correct or incorrect judgments and have neglected their explanations as a principal means of distinguishing one form of reasoning from another.

Judgments and Explanations

Skepticism regarding children's explanations has been based on the assumption that a correct judgment on a given reasoning task reflects a certain logical competence and that an explanation of that judgment reflects the same logical competence as well as a verbal ability to translate it into words. In this view, judgments-plus-explanations provide a less valid indicator of the underlying competence than judgments alone, because the former criterion involves the additional source of error associated with verbal ability as compared with the latter (see Brain-

erd, 1973a, 1977). Against this line of reasoning, we argue that children's judgments alone cannot be assumed necessarily to reflect the same logical competence as their judgments and explanations considered together for the following reasons: (a) The ability to justify or infer a judgment verbally with a certain form of reasoning is itself a logical competence of a certain kind, and (b) this competence is not necessarily reflected in children's judgments considered separately. As indicated in this study, correct judgments can sometimes be inferred equally well with alternate forms of reasoning so that a correct judgment in itself provides no clue about how it was inferred.

Moreover, even a casual reading of Piaget's writings on concrete operations indicates that he classified children's thinking into stages according to the ways in which children justified or inferred their judgments and not merely in terms of their correct or incorrect judgments. Not only could correct judgments sometimes be generated with preoperational as well as with operational thinking (e.g., Piaget, 1941/1952; Piaget & Inhelder, 1941/1974), but operational thinking did not invariably lead to a correct judgment (Piaget, 1974/1978, Chap. 10). As these examples indicate, the operational competencies of interest to Piaget were conceptually grounded in both judgments and explanations, and such competencies must be distinguished from other, purely behavioral competencies that may under some conditions suffice for generating correct judgments. The abandonment of children's explanations as evidence has meant that researchers have deprived themselves of criteria for distinguishing between the two types of competencies and for determining when procedural innovations result in assessing the same or different competencies (cf. Chapman, 1987b).

Once children's forms of reasoning as such (and not merely correct vs. incorrect judgments) are understood as the dependent variable of interest, one can then ask about the antecedent conditions leading children to use one form of reasoning rather than another (cf. Chapman, 1987a; Overton & Newman, 1982). In these terms, the present study provides information on the task-situational conditions influencing children's use of functional versus operational reasoning.

References

- Braine, M. D. S. (1959). The ontogeny of certain logical operations: Piaget's formulation examined by nonverbal methods. *Psychological Monographs: General and Applied*, 73(5, Whole No. 475), 1-43.
- Braine, M. D. S. (1964). Development of a grasp of transitivity of length: A reply to Smedslund. *Child Development*, 35, 799-810.
- Brainerd, C. J. (1973a). Judgments and explanations as criteria for the presence of cognitive structures. *Psychological Bulletin*, 79, 172-179.
- Brainerd, C. J. (1973b). Order of acquisition of transitivity, conservation, and class inclusion of length and weight. *Developmental Psychology*, 8, 105-116.
- Brainerd, C. J. (1973c). The origins of number concepts. *Scientific American*, 228(3), 101-109.
- Brainerd, C. J. (1977). Response criteria in concept development. *Child Development*, 48, 360-366.
- Brainerd, C. J. (1979). *The origins of the number concept*. New York: Praeger.
- Brainerd, C. J., & Kingma, J. (1984). Do children have to remember to reason? A fuzzy theory of transitivity development. *Developmental Review*, 4, 311-377.
- Brainerd, C. J., & Kingma, J. (1985). On the independence of short-

- term memory and working memory in cognitive development. *Cognitive Psychology*, 17, 210–247.
- Breslow, L. (1981). Reevaluation of the literature on the development of transitive inferences. *Psychological Bulletin*, 89, 325–351.
- Bringuier, J.-C. (1980). *Conversations with Jean Piaget*. Chicago: University of Chicago Press.
- Bruner, J. (1983). State of the child. *New York Review*, 30(16), 83–89.
- Bryant, P. (1974). *Perception and understanding in young children*. London: Methuen.
- Bryant, P. E., & Trabasso, T. (1971). Transitive inferences and memory in young children. *Nature*, 232, 456–458.
- Case, R. (1985). *Intellectual development: Birth to adulthood*. Orlando, FL: Academic Press.
- Chalmers, M., & McGonigle, B. (1984). Are children any more logical than monkeys on the five-term series problem? *Journal of Experimental Child Psychology*, 37, 355–377.
- Chapman, M. (1987a). Piaget, attentional capacity, and the functional implications of formal structure. In H. W. Reese (Ed.), *Advances in child development and behavior*, (Vol. 20, pp. 229–334). Orlando, FL: Academic Press.
- Chapman, M. (1987b). Inner processes and outward criteria: Wittgenstein's importance for psychology. In M. Chapman & R. A. Dixon (Eds.), *Meaning and the growth of understanding: Wittgenstein's significance for developmental psychology* (pp. 103–127). Berlin: Springer-Verlag.
- Cohen, D. (1983). *Piaget: Critique and reassessment*. London: Helm.
- DeBoysse-Bardies, B., & O'Regan, K. (1973). What children do in spite of adults' hypotheses. *Nature*, 246, 531–534.
- Fischer, K. W. (1980). A theory of cognitive development: The control and construction of hierarchies of skills. *Psychological Review*, 87, 477–531.
- Gelman, R., & Baillargeon, R. (1983). A review of some Piagetian concepts. In P. H. Mussen (Series Ed.) and J. H. Flavell & E. M. Markham (Vol. Eds.), *Handbook of child psychology* (Vol. 3, pp. 167–230). New York: Wiley.
- Halford, G. S. (1982). *The development of thought*. Hillsdale, NJ: Erlbaum.
- Halford, G. S. (1984). Can young children integrate premises in transitivity and serial order tasks? *Cognitive Psychology*, 16, 65–93.
- Halford, G. S., & Galloway, W. (1977). Children who fail to make transitive inferences can remember comparisons. *Australian Journal of Psychology*, 29, 1–5.
- Halford, G. S., & Kelly, M. E. (1984). On the basis of early transitivity judgments. *Journal of Experimental Child Psychology*, 38, 42–63.
- Kallio, K. (1982). Developmental change on a five-term transitive inference. *Journal of Experimental Child Psychology*, 33, 142–164.
- McGonigle, B., & Chalmers, M. (1977). Are monkeys logical? *Nature (London)*, 267, 694–696.
- Murray, J. P., & Youniss, J. (1968). Achievement of inferential transitivity and its relation to serial ordering. *Child Development*, 39, 1259–1268.
- Overton, W. F., & Newman, J. L. (1982). Cognitive development: A competence-activation/utilization approach. In T. M. Field, A. Huston, H. C. Quay, L. Troll, & G. E. Finley (Eds.), *Review of human development*, (pp. 217–241). New York: Wiley.
- Pascual-Leone, J. (1970). A mathematical model for the transition rate in Piaget's development stages. *Acta Psychologica*, 32, 301–345.
- Perner, J., & Aebi, J. (1985). Feedback-dependent encoding of length series. *British Journal of Development Psychology*, 3, 133–141.
- Perner, J., & Mansbridge, D. G. (1983). Developmental differences in encoding length series. *Child Development*, 54, 710–719.
- Perner, J., Steiner, G., & Stachelin, C. (1981). Mental representation of length and weight series and transitive inferences in young children. *Journal of Experimental Child Psychology*, 31, 177–192.
- Piaget, J. (1952). *The child's conception of number*. London: Routledge & Kegan Paul. (Originally published 1941)
- Piaget, J. (1957). *Logic and psychology*. New York: Basic Books.
- Piaget, J. (1967). *Six psychological studies*. New York: Vintage.
- Piaget, J. (1971). The theory of stages in cognitive development. In D. R. Green (Ed.), *Measurement and Piaget* (pp. 1–11). New York: McGraw-Hill.
- Piaget, J. (1972). *Essai de logique opératoire* [Essay in operatory logic]. Paris: Dunod. (2nd ed. of *Traité de logique*, 1949)
- Piaget, J. (1974). *Understanding causality*. New York: Norton.
- Piaget, J. (1978). *Success and understanding*. London: Routledge & Kegan Paul. (Originally published 1974)
- Piaget, J., Grize, J. B., Szeminska, A., & Vinh Bang (1977). *Epistemology and psychology of functions*. Dordrecht: Reidel. (Originally published 1968)
- Piaget, J., & Inhelder, B. (1974). *The child's construction of quantities*. London: Routledge & Kegan Paul. (Originally published 1941)
- Piaget, J., Inhelder, B., & Szeminska, A. (1960). *The child's conception of geometry*. New York: Harper Torchbooks. (Originally published 1948)
- Riley, C. A. (1976). The representation of comparative relations and the transitive inference task. *Journal of Experimental Child Psychology*, 22, 1–22.
- Riley, C. A., & Trabasso, T. (1974). Comparatives, logical structures, and encoding in a transitive inference task. *Journal of Experimental Child Psychology*, 17, 187–203.
- Russell, J. (1981). Children's memory for the premises in a transitive measurement task assessed by elicited and spontaneous justifications. *Journal of Experimental Child Psychology*, 31, 300–309.
- Smedslund, J. (1963). Development of concrete transitivity of length in children. *Child Development*, 34, 389–405.
- Smedslund, J. (1965). The development of transitivity of length: A comment on Braine's reply. *Child Development*, 36, 577–580.
- Smedslund, J. (1966). Performance on measurement and pseudomeasurement tasks by five- to seven-year-old children. *Scandinavian Journal of Psychology*, 7, 81–92.
- Thayer, E. S., & Collyer, C. E. (1978). The development of transitive inference: A review of recent approaches. *Psychological Bulletin*, 85, 1327–1343.
- Trabasso, T. (1975). Representation, memory, and reasoning: How do we make transitive inferences? In A. D. Pick (Ed.), *Minnesota Symposium on Child Psychology* (Vol. 9, pp. 135–172). Minneapolis: University of Minnesota Press.
- Trabasso, T. (1977). The role of memory as a system in making transitive inferences. In R. V. Kail & J. W. Hagen (Eds.), *Perspectives on the development of memory and cognition* (pp. 333–366). Hillsdale, NJ: Erlbaum.
- Youniss, J., & Furth, H. J. (1973). Reasoning and Piaget. *Nature*, 244, 314–315.
- Youniss, J., & Murray, J. P. (1970). Transitive inference with nontransitive solutions controlled. *Developmental Psychology*, 2, 169–175.

Received July 16, 1986

Revision received August 6, 1987

Accepted December 21, 1987 ■