DOCUMENT RESUME

ED 047 790 24 PS 004 416

AUTHOR Rosenthal, Ted L.; Zimmerman, Barry J.
TITLE Modeling by Exemplification and Instruction in Training Conservation.
INSTITUTION Arizona Univ., Tucson. Arizona Center for Early Childhood Education.
SPONS AGENCY National Center for Educational Research and Development (DHEW/OE), Washington, D.C. Division of Educational Laboratories.
BUREAU NO BR-7-0706
PUB DATE [70]
CONTRACT OEC-3-7-070706-3118
NOTE 32p.

EDRS PRICE EDRS Price MF-$0.65 HC-$3.29
DESCRIPTORS Anglo Americans, *Conservation (Concept), Disadvantaged Youth, Feedback, Grade 1, *Learning Processes, Mexican Americans, Middle Class, *Models, Observation, *Stimulus Generalization, *Training Techniques
IDENTIFIERS Piaget

ABSTRACT

Four experiments (involving a total of 158 children) demonstrated observationally-induced learning effects on multidimensional conservation tasks using brief, single session training. First grade samples of middle class Anglo-American, and economically disadvantaged Mexican-American children displayed gains in vicariously-induced conservation as did a separate group of 4-year-olds. On initial (imitative) task stimuli, modeling increased conservation responses which then generalized to a new set of stimulus items without further intervention. Verbal feedback praising the model's responses did not affect performance. A nonconserving model reduced conservation in initially-conserving children. A nonmodeling instruction procedure failed to modify conservation scores. The model's provision of a rule to explain stimulus equivalence improved performance when criterion response required judged equivalence plus explanation, but not when the response criterion was judged equivalence alone. Children who observed the model conserve without giving an explanation nevertheless increased their correct judgments-plus-rule responses in the imitation phase, thus demonstrating that observational learning can produce inferential thinking. Theoretical implications of the results were discussed. (Author)

ARIZONA CENTER
FOR EARLY CHILDHOOD EDUCATION
College of Education
University of Arizona
1515 East First Street
Tucson, Arizona 85719

MODELING BY EXEMPLIFICATION AND INSTRUCTION IN TRAINING CONSERVATION

By

Ted L. Rosenthal & Barry J. Zimmerman

MODELING BY EXEMPLIFICATION AND INSTRUCTION IN TRAINING CONSERVATION[1]

Ted L. Rosenthal and Barry J. Zimmerman

University of Arizona

Piaget, (Piaget & Inhelder, 1941; Piaget, Inhelder & Szeminska, 1960), in first drawing attention to conservation phenomena, noted that a child's ability to maintain quantitative constancy when stimuli are perceptually transformed increases with age and that conservation proceeds differentially for various stimulus dimensions. Furthermore, he attributes these phenomena to relatively immutable sequences of cognitive development through which, presumably, all children must pass. In support of Piaget, conservation studies have regularly found that children under age 7 do not conserve, that children between ages 7 and 11 conserve on some tasks, and that older children conserve consistently (Inhelder & Piaget, 1958.) Although questioning that such general results really confirm Piaget's notion of invariant stages of intellectual structure, Kingsley and Hall (1967) acknowledge that considerable normative data also support age-related effects for divergent stimulus dimensions, namely that conservation of substance or mass typically precedes conservation of weight or length which, in turn, occur before conservation of volume and number. Early attempts to create "precocious" conservation by training young children usually failed, and reviewing that literature Flavell (1963) concluded that although most of the methods attempted appeared sound and reasonable "most of them have had remarkably little success in producing cognitive change." Subsequent research using several "discovery" methods (Mermelstein & Meyer, 1969) also failed to create changes, even with large samples of children.

However, recent behaviorally-oriented studies have demonstrated that training can influence conservation behavior. Following Gagne's (1965) task-analysis approach, Kingsley and Hall (1967) trained children on a graded series of subtasks using somewhat variable techniques, flexibly tailored to individual youngsters. Their subjects displayed increased weight conservation that generalized to conservation of substance. In another study of sequentially-graded practice on component skills, Ruthenberg and Orost (1969) used a variety of training procedures for number conservation, including the aid of slightly older peers as "assistant teachers" to discuss with (but not demonstrate for) subjects the bases of correct response. Rothenberg and Orost were thus able to instate number conservation in kindergarten children who then generalized conservation to discontinuous quantity; these effects were maintained upon retesting after a three-month delay. In a carefully-controlled experiment, Gelman (1969) emphasized discrimination learning, rather than skill practice, by teaching 5-year-olds to attend to quantitative relations and to exclude the other stimulus attributes of practice problems. After extensive training, she found that exposure to relevant examples, plus feeding back knowledge of results of his response to the child, together did effectively elicit conservation of length and number. A group of children given the same discrimination training without feedback showed much poorer performance. Consistent with her hypotheses, Gelman further found that stimulus features defined as irrelevant created conservation errors when children failed to ignore such cues. In all the recent and successful studies described above, rather exhaustive direct task- or discrimination-practice has been needed to modify relatively delimited subclasses of conservation; this point is illustrated by the stimulus dimensions italicized in the foregoing paragraph.

Working from somewhat different theoretical emphases, students of social learning have demonstrated that a wide variety of affective, motor, and self-regulatory behavior can be rapidly modified by observation of a model (e.g. Bandura, 1969a, 1969b, 1971). However, rather little research in the social learning tradition has explicitly concerned the modification of abstract, "higher-level", conceptual behavior. In a previous report, Rosenthal, Zimmerman, and Durning (in press) have shown that by observation children rapidly learned a model's diverse question-formulations to a set of stimulus pictures and generalized these styles of inquiry to new stimuli without further training. In four variations, the model exemplified alternative question-types to the same pictures. For separate groups of children her questions were based on the nominal and physical properties of the objects depicted, on functional uses to which objects might be put, on causal relationships involving the objects, or on judgments of value and preference regarding the objects. The model never gave children any verbal rubric to summarize the criteria governing her question-categories; nevertheless, in all variations the children proved able to <u>induce</u> the categorical features of the model's styles of inquiry as shown by imitative reproduction of these features to the original stimuli and, without further training, transfer of these features to a novel set of stimulus pictures. These vicarious effects upon abstract, cognitive behavior were obtained with relatively mature, sixth-grade samples; it thus seemed desirable to extend the study of social learning influences upon concept-formation to younger children and to a cognitive task of traditional interest.

The present experiments tested the effects of observing a model on children's conservation judgments. To avoid confining results to any particularized subclass of conservation (e.g. number versus weight versus area), an adaptation of the Goldschmid and Bentler (1968a, 1968b) Concept Assessment Kit was employed. This instrument, which is available in alternate forms, provides a generalized measure of conservation across a number of stimulus instances, such as shape properties in two- and three-dimensional space, and constancy of volume for both continuous and discontinuous quantitites. Goldschmid and Bentler (1968a) describe their scales as representing "measures of a general concept of conservation — like a general factor of factor analysis", and provide strong psychometric evidence in support of their interpretation. In the present research, after a baseline phase using the first scale (Form A), experimental subjects observed the model's performance; in certain conditions, (below) an explanatory rule accompanied her conservation judgments. Form A was then readministered to all children and, to assess generalization effects, a new but closely-comparable scale (Form B) was subsequently presented without further training to all children. Consider sample items in which equal amounts of water (or parched corn) are first presented to the child in two identical containers (e.g. glasses). Then, the contents of one vessel are transferred to a container of different shape (e.g. a broader or a tall, narrow glass). To be credited with "conserving", the child must recognize that despite perceptual transformation of one member, the paired stimulus-members remain quantitatively equal (i.e. contain equal amounts of water or corn).

Piaget's definition of conservation (1952) required that the child both judge the stimulus members as equal, despite one having been perceptually transformed, and that the child offer some logical reason for his equivalence judgment. Although providing separate scores for equivalence judgments and for explanations, following Piaget, Goldschmid and Bentler (1968a, 1968b) have chosen to require that both judgment and reason be correct in order to credit an item. Since no *a priori* basis exists for assuming that correct apprehension of a concept is necessarily accompanied by the capacity to give a plausible justification for one's answer, in the present studies both the response-criterion of correct judgment alone, and the criterion of correct judgment *plus* correct explanation, were investigated and were separately analyzed.

In overview, Experiment I factorially compared observationally-instated conservation as a function of rule-provision versus no explanation by the model, and verbal praise versus no feedback administered to the model's responses by the experimenter. Experiment II studied the effect of a *non*-conserving model upon the judgments of children who, initially, had given some evidence of ability to conserve. Experiment III compared observing a model demonstrate conservation judgments versus a non-modeling procedure in which the experimenter verbally informed children that the transformed stimuli were equal, using Mexican-American youngsters from a barrio area. Experiment IV studied model-created conservation in a group of very young, pre-school, children. In all studies, the Form A items were presented in baseline, and again in the imitation phase, and then the new, Form B items were introduced to assess generalization phase effects.

Experiment I

METHOD

Subjects. From two schools in Tucson, 50 boy and 50 girl first-graders with baseline scores of four or less were randomly drawn. (If a child, however, scored five or six in baseline, he was reserved for Experiment II, below, for non-conservation modeling; 17 children were thus isolated.) Equal numbers of boys and girls were randomly assigned to the experimental conditions or the control group. Nearly all youngsters were Anglo-American and from middle-class homes. Their ages ranged from 5.9 to 6.8 years with a median age of 6.3; all subjects were in their first semester of public school.

Procedure. The Goldschmid and Bentler (1968a) measure of generalized conservation, which is available in closely-parallel, six-item Forms (A and B) was used. This instrument examines the spatial, substance, weight, number, and continuous- and discontinuous-quantity dimensions of conservation, with different item-stimuli provided for Forms A and B. On all items, the child is presented with a pair of visually-equivalent stimuli; next, one member of the pair is perceptually transformed by the experimenter (e.g. one of two equal balls of clay is rolled into a cylinder), and then the child is asked if the objects still are equivalent. Norms are provided for conservation defined by the judgment of equivalence alone, and by the equivalence-judgment plus an explanation for equivalence (e.g. "if you rolled the clay back into a ball, the balls would still be the same.") Before transforming stimuli, Goldschmid and Bentler (1968a) obtained the child's agreement that the members of stimulus pairs were the same. If the child failed to agree, the stimuli were again distributed until the subject accepted their equality. In the present procedures, before transforming a stimulus the child was told that paired stimuli (e.g. the two balls of clay) were equal and redistributions were made only if, (as did some five youngsters), the child disagreed.

Form A was given to all children in a baseline phase and, following presentation of modeling (or control) treatments, it was immediately readministered in an imitation phase to assess vicariously-created (and control) changes. Directly after, in a generalization phase, Form B was introduced to all children without further training to determine the transfer of conservation behavior to new stimulus items.

Each child was taken from class to a separate room where he worked with the adult, male experimenter and the adult, female model. Each child was seated at a low table adjoining a bookcase that shielded all testing materials from his view, except when they were in use. During the experiment, the model recorded the child's responses on a record sheet; she thus was present during data collection for all children. Before baseline, the child was asked to count 16 red barns to establish that his quantitative skills were adequate to numerically discriminate forthcoming items (e.g. when 16 blocks were transformed from a square to a triangular display). All children met this counting criterion, after which the first item of Form A was presented. Care was taken during baseline to prevent the child from watching the return of transformed stimuli to their original appearance (which would have given children cues regarding the reversability of the perceptual changes), by having the experimenter make these re-transformations behind the bookcase, out of sight.

After baseline, experimental subjects were instructed as follows: "Now let's give this lady a chance to play the games. I want you to watch and listen carefully, and you will have a chance to play the games again later." The model then performed and, subsequently, the experimenter retransformed the stimuli out of sight of the child and then introduced the imitation phase by saying: "Now, let's play these games again." To assess "spontaneous" changes without training, the italicized instructions were given to control subjects directly after baseline. Form A was then readministered and next Form B was presented to all subjects with the following introduction: "Here are some little bit different (sic.) games for you to play."

Treatments. Equal numbers of boys and girls were randomly assigned to each of four experimental groups or to the no-model, control condition already described. All experimental subjects observed the model give an equivalence judgment to each item of Form A, using one of three verbal formats to avoid repetition of precise wording: "There's just as much here as is there."; "They're both the same."; or "There is the same amount here as is there."

For half the experimental subjects, the experimenter asked the model to give an explanation for her judgment of equivalence. The model then supplied the following rule: "Because they were the same in the first place." Rule-provision was omitted for the remaining children. As a second variable, feedback was studied by having feedback groups observe the experimenter offer verbal reinforcement to the model ("that's right", "that's good", or "correct") after each of her responses; verbal praise to the model was omitted in no-feedback conditions. The four groups formed by combining rule X feedback treatments each contained 10 boys and 10 girls. As noted earlier, analyses were separately computed for conservation defined by just equivalence judgments (judgments-only), and for the equivalence judgments plus explanation (judgments-plus-rule) criterion.

RESULTS

Judgments-Only. Before presenting the major results, it is necessary to consider the performance of the no-model control group. For these uninfluenced control subjects, the correlation between baseline and imitation phase scores on Form A equaled .77, and the r between baseline and the Form B generalization items equalled .80. It can thus be seen that the conservation score measures maintained substantial test-retest and alternate-form reliabilities despite the minor alterations of procedure required by the present design.

A repeated-measures analysis of variance across baseline, imitation, and generalization phases revealed "practice effect" increases that approached significance ($F = 3.04$; $df = 2/38$; $p < .06$)[2]. When the combined experimentals were compared with the controls across phases, a highly significant trials effect ($F = 176.68$; $df = 2/196$; $p < .001$) was obtained, indicating systematic score increases from baseline. Furthermore, significant groups ($\underline{F}$ = 24.80; $\underline{df}$ = 1/98; $\underline{p} < .001$), and groups X trials interaction ($\underline{F}$ = 34.34; $\underline{df}$ = 2/196; $\underline{p} < .001$) effects revealed that, aggregately, the experimental subjects had far surpassed the score-increases of the control group; thus, the modeling variations had increased conservation scores over and above any practice effects based on sheer exposure to item stimuli. Dunnet's test (Kirk, 1968) was applied to the several components of the interaction term and disclosed that, although experimental and control subjects did not differ significantly at baseline, the experimentals significantly surpassed control performance at the imitation and the generalization phases (both $\underline{p}$'s $< .01$). The score means at each phase for the control and separate experimental groups are presented in Table 1.

---

insert Table 1 about here

The main 2 X 2 X 3 factorial analysis (comparing rule and feedback treatments across phases) only yielded a significant effect for trials ($\underline{F}$ = 176.67; $\underline{df}$ = 2/152; $\underline{p} < .001$) indicating that scores increased from baseline across phases for all modeling groups, with no other significant main effects or interactions. Thus, neither approving feedback from the experimenter directed toward the model's responses, nor provision of the explanatory rule by the model, influenced conservation behavior as defined by the criterion of equivalence-judgments only. *Post hoc* comparisons with Tukey's HSD test (Kirk, 1968) disclosed that imitation and generalization phase means were each significantly higher than baseline conservation (both p's $<$ .01); performance did not decrease significantly from imitation to generalization phases.

Judgments-plus-Rule. Again, before presenting the major results, the performance of the no-model control group requires comment. It is of interest that control subjects obtained a correlation .79 between baseline and imitation scores with the Form A itmes, and an $r = .74$ between baseline and generalization scores using the Form B stimuli. Thus, the conservation measures again displayed substantial test-retest and alternate-form reliabilities despite the minor procedural alterations imposed. Furthermore, the correlations between the judgments-only and the judgments-plus-rule criteria for baseline, imitation, and generalization phases equaled .79, .85, and .83, respectively. These coefficients confirm that the alternative response-criteria did reflect empirically similar definitions of conservation.

When the control group's scores were assessed across phases by a single-group, repeated-measures analysis of variance, a trials effect approaching significane ($F = 2.98$; $df = 2/38$; $p < .07$) was observed. However, when the controls were compared with the combined experimental subjects an overall trials effect ($F = 62.78$; $df = 2/196$; $p < .001$), a significant groups effect ($F = 8.19$; $df = 1/98$; $p < .01$), and significant groups X trials interaction ($F = 11.37$; $df = 2/196$; $p < .001$) were found, demonstrating superior conservation by the experimental subjects who had observed the model perform. Subsequent post hoc comparisons with Dunnet's test revealed that although experimentals and controls had not differed at baseline, the experimental subjects significantly surpassed the control group in the imitation and the generalization phases (both p's $< .01$). The score means by phase for the controls and the separate experimental variations are presented in Table 1.

The general pattern of the major findings, (combining feedback and no feedback groups for the rule and no-rule modeling conditions separately), may be better ascertained from Figure 1 which presents the means in baseline, imitation, and generalization phases for the rule and no-rule modeling groups, and for the no-model controls.

---

insert Figure 1 about here

The main 2 X 2 X 3 factorial analysis among modeling conditions revealed a significant trials effect ($F = 66.32$; $df = 2/152$; $p < .001$), a significant main effect for rule-provision ($F = 7.25$; $df = 1/76$; $p < .01$), and a significant interaction term for rule-provision X trials ($F = 6.76$; $df = 2/152$; $p < .01$). The experimenter giving favorable feedback to the model's responses failed to influence conservation ($F = 0.87$; $df = 1/76$; NS) or to interact in any fashion with rule-provision or change across trials (largest $F = 2.11$; $df = 2/152$; NS). Tukey tests revealed that although rule and no-rule modeling groups did not differ at baseline, each group exceeded its own baseline scores both in the imitation and the generalization phases (all p's $< .01$). Furthermore, by the scoring criterion requiring judgment plus explanation, children who observed the model give the rule with her equivalence-judgments surpassed the no rule treatment in the imitation and the generalization phases (both p's $< .01$).

Scrutiny of Figure 1 suggests that the no-rule modeling condition had improved on the judgments-plus-rule criterion from observing a model who had <u>not</u> provided the rule with her equivalence judgments. To confirm that the no-rule group had induced scoreable explanations beyond any sheer-exposure "practice effects", the no-rule modeling condition (combining feedback variations) was compared across phases with the control group. An overall analysis of variance revealed a significant across-trials effect (F = 19.71; df = 2/116; $p < .001$) and a borderline between-groups effect (F = 2.95; df = 1/58; $p < .10$), in the direction of superior conservation by the no-rule modeling subjects. In addition, a significant groups X trials interaction term (F = 5.51; df = 2/116; $p < .01$) was found. Further analysis of this interaction by Dunnet's test disclosed that although the no-rule experimentals did not significantly differ from the controls in baseline or generalization phases, they did significantly surpass the controls in the imitation phase ($p < .01$). Thus, without rule-provision, observing the model give equivalence judgments led children to increase their production of correct explanations in the imitation phase.

It seemed of interest to determine whether observing a model fail to give conservation judgments would reduce conservation in children who, originally, had shown evidence of being able to conserve. Accordingly, Experiment II was conducted.

Experiment II

METHOD

Subjects and Procedure. The subjects were that subset of 17 children who had conserved on at least five of the six baseline items. The 10 boys and 7 girls so selected were, in age and family socioeconomic background, typical of the larger sample earlier described. The present procedures closely paralleled those of Experiment I, and were executed by the same adult experimenter and model. Presentation of the barn counting task and of baseline items was conducted identically as before. If, however, a child judged the transformed stimuli (by the judgments-only response criterion) to be equivalent on at least five baseline items, he was for training assigned to observe the model demonstrate non-conservation judgments.

To illustrate the model's non-conserving judgments, consider an item in which two equal balls of Play Doh were first presented as a pair of stimulus members; then, one of the balls (a) was flattened into a "pancake" by the experimenter while the other member (b) remained untransformed. After such transformations, for each item the manual (Goldschmid and Bentler, 1968a) designated one member of each stimulus-pair as a, and the other member as b. On the first item, the model selected the a member by pointing at it and stating: "That one has more." The model continued in like fashion thereafter, alternating between a and b members on successive items. The model never gave any explanation for judging stimulus members to be unequal, and the experimenter neither praised nor criticized the model's responses. Thus, for the child observers, the model's demonstration created conditions essentially opposite in concept, but analogous in method, to those of the no rule-provision, no-feedback group of Experiment I. Procedures identical with those of Experiment I were instituted after modeling, in the imitation and the generalization phases; the only noteworthy difference in "staging" the two studies involved the reversed (non-conservation) direction of the model's judgments for the children with high baseline conservation scores.

RESULTS

Judgments-Only. A single-group, repeated-measures analysis of variance revealed that observing the non-conserving model significantly reduced conservation judgments across trials (F = 13.53; df = 2/32; $p < .001$). Subsequent Tukey tests disclosed that conservation during imitation (M = 2.71) and generalization (M = 3.35) phases were each significantly lower (both p's $< .01$) than were the children's baseline scores (M = 5.35).

Judgments-plus-Rule. It will be recalled that the non-conserving model was never asked to give reasons for her judgments of inequality between paired stimulus-members. Consequently, a significant decrease from baseline in children's production of explanations which favored equivalence between stimulus-members would give further evidence that observing the model's judgments alone could affect, inferentially, the child's production of verbal reasons for his judgment responses. The overall analysis of variance revealed significantly reduced conserving across trials (F = 3.58; df = 2/32; $p < .04$). By Tukey tests, only the mean for the imitation phase (M = 1.94) was itself significantly lower ($p < .05$) than the baseline mean of 3.24; although the generalization scores decreased (M = 2.53), they did not differ significantly from baseline response. Nevertheless, these data provide some evidence of reverse inference, i.e. that after observing the model, the child's original reasons did not remain viable explanations of his altered judgments and hence were used less. Anecdotal evidence further supported this conclusion: Typically, the scoreable reasons given during baseline were dropped after observing the model, and one of varied non-conservation rationales was substituted, e.g. "because it's bigger (or heavier, taller, longer, etc.)."

Experiment III was conducted to study a group of economically-disadvantaged children and to compare the information-transmitting effectiveness of modeling with a non-modeling, verbal instructions technique.

Experiment III

METHOD

Subjects. From the first grade of an elementary school located in an economically depressed area of Tucson, a sample of 28 Mexican-American children from economically-disadvantaged homes was randomly drawn. The school was situated in a barrio (ghetto) community and was receiving Federal support under Title I of the Elementary and Secondary Education Act. The children ranged in age from 6.1 to 7.0 years with a median age of 6.3 years. All children were in their first semester of public school and Spanish was the original language spoken in the home of each child.

Procedure. Seven boys and seven girls were randomly assigned to each of two experimental treatments. The modeling treatment was the same as the with rule-provision, no feedback condition described in Experiment I. As a prototype of conventional pedagogical practice, a non-modeling instructions variation was also studied. The instructions and modeling treatments received identical baseline procedures. After baseline, the experimenter presented the instructions group with the items in sequence such that, for each item, one member of the stimulus pair was already transformed; e.g. a clay ball and a clay cylinder were displayed. The experimenter told the subject that the stimulus-members were equivalent, giving the rule for each item (e.g. "There is just as much wood here as there is here because they both had the same amount in the first place.") Consequently, instructions group subjects heard both the judgment of stimulus equality and the reasons for this equivalence. The two treatments only differed by the modeling group watching the experimenter transform one member of each stimulus-pair and hearing the model verbalize her answers in response to the experimenter's questions whereas, in the instructions treatment, the children were shown stimulus-pairs with one member already transformed, and heard the experimenter paraphrase the judgment of equality and the explanation for stimulus equivalence, but without this information being cast in the question-and-answer format given to the modeling treatment. After training through modeling or instructions, both groups received identical procedures in the imitation phase re-test, and, without any further tutelage, in the generalization phase; these procedures were earlier described in Experiment I.

RESULTS

Judgments-Only. A 2 (treatments) X 3 (phases) analysis of variance revealed significant increases across trials ($F = 23.24$; $df = 2/52$; $p < .001$), a significant between-groups effect in favor of the modeling treatment ($F = 23.46$; $df = 1/26$; $p < .001$), and significant groups X trials interaction ($F = 9.84$; $df = 2/52$; $p < .001$). Tukey tests disclosed that, at baseline, the instructions (M = 0.00) and modeling (M = 0.36) groups did not differ. The modeling group (M = 4.50) surpassed the instructions group (M = 1.00) to a significant extent in the imitation phase ($p < .01$); this superiority of modeling (M = 3.93) over instructions (M = 0.53) was maintained in the generalization phase ($p < .01$). Inspection of the data suggested that not only had the modeling surpassed the instructions group, but that the instructional procedure had failed to create any stable changes. To evaluate this impression, the data for each group relative to its own baseline were separately compared across phases by Tukey tests. The modeling group continued to display strong learning effects from baseline to the imitation, and to the generalization phases (both p's $< .01$); there was no significant decline in performance from imitation to generalization. In contrast, the instructions group gave no evidence of having improved from baseline across phases.

Judgments-plus-Rule. The main 2 X 3 analysis again revealed significant effects for trials ($F = 12.47$; $df = 2/52$; $p < .001$), for groups ($F = 7.13$; $df = 1/26$; $p = .01$), and for groups X trials interaction ($F = 6.71$; $df = 2/52$; $p < .01$). Tukey tests again confirmed that the instructions and modeling groups (both M's = 0.00) did not differ in baseline. The modeling (M = 2.79) group significantly surpassed the instructions group (M = 0.43) during imitation ($p < .01$); the continued superiority of modeling (M = 2.79) over instructions (M = 0.43) procedures remained significant ($p < .01$) in the generalization phase; indeed, the means for each group remained unchanged between imitation and generalization. It was again germane to consider the alternative training treatments separately, in relation to their own baseline scores. Tukey tests confirmed that modeling subjects improved from baseline to imitation and to generalization phases (both p's $< .01$). In contrast, the instructional procedure did not create any performance increase over baseline.

Since Piaget has contended that children below age 7 do not conserve (Piaget & Inhelder, 1958), it seemed important to study a group of very young children in Experiment IV.

Experiment IV

METHOD

Subjects and Procedure. From the small group enrolled in the University of Arizona Nursery Preschool, 7 boys and 6 girls were randomly chosen. The children came from primarily middle class homes of both Anglo- and Mexican-American parentage. The children ranged in age from 4.2 years to 4.9 years, with a median age of 4.6 years.

Pilot testing revealed that the observational training procedures used in the previous experiments appeared to unduly tax youngsters of this age. In particular, the uninterrupted modeling format appeared to exceed the children's capacity to observe with sustained attention. Consequently, an alternation format adapted from Bandura and Harris' (1966) methodology was introduced. In this technique, during training the model would perform on an item using the procedures of the no rule-provision, no feedback condition of Experiment I. Then, the experimenter would return the transformed stimulus-member to its original state (e.g. from a cylinder back to a ball) out of sight of the child and immediately presented the item to the subject. Thus, the model and child alternated in responding sequentially to the six items of Form A. Subsequently, the generalization phase was, as before, introduced without further training and the six Form B items were presented in series with no intervention by the model. To reduce running time, the pre-baseline barn counting task was eliminated; otherwise, the experimental procedures were identical with those specified in Experiment I.

RESULTS

Judgments-Only. Analysis of variance revealed significant increases from baseline across phases (F = 8.28; df = 2/24; p = .002). Tukey comparisons revealed that from no conservation in baseline (M = 0.00), the children increased their judgments of equivalence in the imitation (M = 2.08) and the generalization (M = 1.77) phases to a significant extent (both p's $<$ .01).

Judgments-plus-Rule. The youngsters failed to verbalize scoreable explanations from observing the model perform. Anecdotal evidence suggested that limitation in the verbal repertoires of very young children had influenced this result: For example, the children offered nonscoreable reasons such as "That one was a block and that one is.", and "That one was a ball." These statements appeared informative by the children's standards of social communication, and they would peer incredulously at the experimenter when asked to "Tell me more.", simply repeating their original cryptic revelations. It thus appeared, as seen in context, that the linguistic components requisite for a formally adequate explanation were beyond the verbal repertoires readily accessible to these four-year-olds.

DISCUSSION

Taken together, the present experiments demonstrated rapid and substantial vicariously-produced changes in children's conservation behavior. The model's providing a verbal rationale for her judgments influenced the response-criterion that required verbal explanation plus judgment of stimulus equivalence, but did not improve performance on the judgments-only response criterion. The experimenter giving approving feedback to each of the model's responses failed to affect either the judgments-only or the judgments-plus-rule conservation measures. Experiment III revealed that exemplification, (such that the model's judgments and the physical transformations of stimuli were sequentially associated), was itself a crucial component of conservation training. In this study, Mexican-American barrio children heard equivalence judgments and reasons given for all items, and observed the same terminal arrangements of conservation stimuli. The instructions group who were shown the already-transformed stimuli, and were told by the experimenter that stimulus-members were equivalent because "they both had the same amount in the first place", utterly failed to increase their conservation responses over baseline. In contrast, the exemplification group which observed the model give equivalence judgments and explanations in conjunction with witnessing the experimenter transform the item stimuli gave evidence of significant and substantial increases in conservation on both response-criteria and in both the imitation and generalization phases. One should note that the instructions (non-exemplification) group received training procedures reminiscent of the other studies, earlier described, which eventually did train conservation (e.g. Rothenberg & Orost, 1969). The present superiority of the model-exemplification group suggests that, relative to other behavioral techniques, social learning methods may have special merit for rapidly modifying complex, cognitive behavior in youngsters not from English-language backgrounds. Thus, unlike the instructions technique, a single-session modeling procedure was able to instate significant conservation increases. This disparity in efficiency may help clarify the reasons that the successful, but non-modeling. behavioral techniques considered earlier required such lengthy practice to influence conservation. The results point toward the potential utility of social learning methods for systematic teaching of economically- or linguistically-marginal children by film or television presentation of modeling sequences.

It might have been desirable to study a training procedure in which the model illustrated appropriate responses to transformations rendering paired stimulus-members <u>unequal</u> in their quantitative properties. In such a method, the child could observe both stimulus equivalence and divergence, and a judgment of equality would not always be credited as evidence of conservation. To implement these considerations would have forced wide departure from the standard conditions of administration specified by Goldschmid and Bentler (1968a) and, by the addition of further items, would have excessively fatigued the children, who were already taxed by the duration of the present procedures. However, the vicarious effects on information-seeking earlier described (Rosenthal, Zimmerman, & Durning, in press), under conditions in which a child was free to attain scores via any of myriad responses, demonstrated the induction of rule-governed behavior from observation of widely-diverse instances. Thus, social-learning research using highly varied stimulus and response specifications has obtained independent results in essential agreement with the present findings. It appears of importance that social-learning procedures have now been shown to be effective influences upon inferential thinking. By a variety of comparisons, children not only adopted the models' abstract paradigms, but spontaneously extended them to novel stimulus instances, and displayed qualification of their conceptual responses to accord with the constraints implied by the models' rubrics.

One may anticipate that questions will be raised regarding the extent to which children "really" learned to abstract quantitative invariance between stimulus-members, or "really" just learned to emit verbal judgments of stimulus equivalence. No single demonstration is likely to wholly satisfy such reservations, but a variety of data support the interpretation that subjects were inducing abstract relationships: First, Goldschmid and Bentler (1968a) reported a correlation of +.90 between the judgments-only and the judgments-plus-rule criteria; thus, except when salient experimental operations favor one response variant over the other, an increase in either measure should prove an excellent predictor of increase in the alternative criterion. Second, the model, no rule-provision group of Experiment I displayed a significant increase over their baseline scores on the judgments-plus-explanation measures, and also surpassed the no-model controls by this criterion; thus, without directly having observed reasons for equivalence given, the children seemed able to induce explanations from the model's judgments alone. Third, observation of a non-conserving model's judgments alone (no explanations provided) significantly reduced below baseline frequency the subsequent production of reasons for stimulus equivalence which the children of Experiment II had originally given. If, in the two foregoing instances, the children had merely been "parroting" the model's verbal judgments, it would seem implausible for their explanations to covary, upward and downward respectively, with the import of the model's statements. Fourth, in Experiment III, both groups of children heard similar content, i.e. verbalizations of stimulus equivalence and reasons. Unlike the model-exemplification group, the no-model instructions group failed to improve from baseline. Had the major findings resulted from a mere copying of verbalizations, the instructions group would be expected to show some evidence of increased conservation: If the child were just "copying" or "complying", he should have emulated the judgments and reasons which the experimenter provided.

Furthermore, it is implicit in Piaget's conservation requirement of equivalence judgment plus explanation that offering a reason is evidence of greater understanding than is judged equivalence alone. The inference that ability to correctly verbalize stimulus relationships (i.e. give an explanation) necessarily presupposes "real" comprehension has been questioned by current research. In a series of studies on inferential problem-solving, Kendler and Kendler (1967) trained youngsters to discriminate between objects differentially related to goal-attainment. Two groups of children learned to label these objects, whereas a third group did not. The results indicated that access to overt labels might assist with, interfere with, or fail to influence problem-solution. In the present context, these data caution that a conservation criterion which includes a verbal reason cannot simply be assumed to assure better grasp of abstract concepts than a judgments-only criterion. Similarly, in a study of the relationship between performance on a transposition task, and children's ability to correctly verbalize the stimulus arrangements, Morris (1970) found correct transposition was not associated with verbalization in any simple fashion. Morris concluded that "the presence of overt verbal labels may either facilitate or interfere with young children's problem-solving behavior." and that "children often have difficulty in maintaining verbal concepts and subsequently integrating these concepts with overt responses." Somewhat reminiscent of the present results, Morris also found that feedback from experimenter to child concerning the correctness of the child's verbalizations of stimulus relationships neither improved transposition responses nor post-transposition verbalizations.

The present procedures differed from other experimental attempts to train conservation by using composite scores, based on items representing several diverse dimensions of conservation. The more usual strategy has been to select a single stimulus-dimension for training, and perhaps to study generalization to another dimension. Therefore, the propriety of combining perceptually-discrepant dimensions into an aggregate score may warrant comment. It is important to recognize that Goldschmid and Bentler's (1968a, 1968b) rationale for their generalized conservation measure is mainly empirical: they found very substantial internal consistency coefficients (both for judgments-only and judgments-plus-rule scoring criteria) among the items representing their several stimulus dimensions; they further reported exptensive evidence for an essential homogeneity among the subvarieties of conservation comprised by their test. Since some item-difficulties found by Goldschmid and Bentler departed from the hypothetical developmental sequence of conservation-attainment summarized by Kingsley and Hall (1967), research studies designed to promote "precocious" conservation might best use composite measures. Such an approach would forestall doubts concerning whether procedures effective in training conservation on a presumably "early-maturing" dimension (e.g. mass) would also succeed with a presumably "late-maturing" dimension (e.g. volume). It would seem desirable for attempts to demonstrate training effects to avoid controversy regarding the differential rates of "ripening" of the several conservation dimensions, until these have been better explicated experimentally.

That verbal praise addressed to the model's responses by the experimenter failed to influence conservation scores suggests that the main results cannot be readily attributed to compliance effects: If social-influence processes were eliciting "superficial" compliance (e.g. Kelman, 1961), or if conservation responses were merely efforts to meet situational demands for socially-desirable behavior, it would necessarily be expected that the groups observing feedback ("correct", "that's right") to the model should have faced the clearest demand characteristics and the strongest social-influence pressures. In fact, observation of approving feedback had no effect upon conservation performance. It seems more probable that when an adult experimenter and model in a school setting provide students with a test-like task, that generalization from recurrent experiences with teachers, and tutoring parents, would create strong pressures to excel and to please the adults independent of the child's assignment to the experimental or control variations presently investigated.

References

Bandura, A. Principles of behavior modification. New York: Holt, Rinehart & Winston, 1969a.

Bandura, A. Social-learning theory of identificatory processes. In D. A. Goslin (Ed.) Handbook of socialization theory and research. Chicago: Rand McNally, 1969b.

Bandura, A. Theories of modeling. New York: Atherton Press, 1971.

Bandura, A., & Harris, M. B. Modification of syntactic style. Journal of Experimental Child Psychology, 1966, 4, 341-352.

Flavell, J. H. The developmental psychology of Jean Piaget. Princeton, N. J.: Van Nostrand, 1963.

Gagne, R. M. The conditions of learning. New York: Holt, Rinehart & Winston, 1965.

Gelman, R. Conservation acquisition: a problem of learning to attend to relevant attributes. Journal of Experimental Child Psychology, 1969, 7, 167-187.

Goldschmid, M., & Bentler, P. M. Manual: concept assessment kit, conservation. San Diego: Educational and Industrial Testing Service, 1968a.

Goldschmid, M., & Bentler, P. M. Dimensions and measurement of conservation. Child Development, 1968b, 39, 787-802.

Inhelder, B., & Piaget, J. The growth of logical thinking from childhood to adolescence. New York: Basic Books, 1958.

Kelman, H. C. Processes of opinion change. Public Opinion Quarterly, 1961, 25, 57-78.

Kendler, T. S., & Kendler, H. H. Experimental analysis of inferential behavior in children. In L. P. Lipsett & C. C. Spiker (Eds.) Advances in child development and behavior. Vol. 3. New York: Academic Press, 1967.

Kingsley, R. C., & Hall, V. C. Training conservation through the use of learning sets. Child Development, 1967, 38, 1111-1126.

Kirk, R. E. Experimental design: procedures for the behavioral sciences. Belmont, Calif.: Brooks/Cole, 1968.

Mermelstein, E., & Meyer, E. Conservation training techniques and their effects on different populations. Child Development, 1969, 40, 471-490.

Morris, L. A. Relational responding to a transposition task as a function of relevant verbalizations and feedback. Unpublished doctoral dissertation, University of Arizona, 1970.

Piaget, J. The child's conception of number. New York: Humanities, 1952.

Piaget, J., & Inhelder, B. Le développement des quantités chez l'enfant. Neuchâtel: Delachaux et Niestlé, 1941.

Piaget, J., Inhelder, B., & Szeminska, A. The child's conception of geometry. New York: Basic Books, 1960.

Rosenthal, T. L., Zimmerman, B. J., & Durning, K. Observationally-Induced changes in children's interrogative classes. Journal of Personality and Social Psychology, in press, 1970.

Rothenberg, B. B., & Orost, J. H. The training of conservation of numbers in young children. Child Development, 1969, 40, 707-726.

Footnotes

1. These studies were supported by the Arizona Center for Early Childhood Education as a Subcontractor under the National Program on Early Childhood Education of the Central Midwestern Regional Educational Laboratory, a private non-profit corporation supported in part as a regional educational laboratory by funds from the United States Office of Education, Department of Health, Education, and Welfare. The opinions expressed in this publication do not necessarily reflect the position or policy of the Office of Education, and no official endorsement by the Office of Education should be inferred. We wish to acknowledge the generous cooperation of Principals E. B. Appleman, R. W. Roda, and R. B. Stanley, of their teachers and the administration of Tucson School District 1, and of Mrs. Jean R. Dees and her staff at the University of Arizona Nursery Pre-school. M. Ferg, M. Hurt, Jr., and B. Wexler provided valuable assistance with diverse aspects of this research. We are grateful to Albert Bandura and to Glenn M. White for their helpful suggestions regarding the manuscript.

2. All tests of significance reported in this paper were based on two-tailed probability estimates.

Table 1

Mean Conservation Responses by Phase for Experimental and Control Groups on Judgments-Only, and Judgments-plus-Rule Response Criteria

Group	Response Criterion: Judgments-Only, Phase: Base-line	Imita-tion	General-ization	Judgments-plus-Rule, Phase: Base-line	Imita-tion	General-ization
No rule no feedback	0.50	4.90	4.50	0.25	2.45	2.10
No rule with feedback	0.50	4.90	4.30	0.35	1.60	1.55
Rule no feedback	0.50	4.20	3.90	0.20	4.10	3.35
Rule with feedback	0.85	4.20	4.10	0.45	3.35	3.00
Control	0.65	0.90	1.20	0.45	0.65	0.95

Note. — Control subjects observed no model but were given retest and generalization stimuli to assess "spontaneous" changes.

Figure Caption

Fig. 1. Mean correct judgments-plus-rule responses by phase for no-rule and with-rule modeling groups (combining feedback variations), and for the control group.