Exploring Stimulus Equivalence Formation in Infants

Martha Peláez

Jacob L. Gewirtz

Aida Sanchez

Nadia M. Mahabir

Florida International University

A systematic replication of the Devany, Hayes and Nelson (1986) and the Augustson and Dougher (1992) studies was conducted with infants to explore the relation of stimulus equivalence to language development. 9 normal infants, age 21 to 25 months participated in 5 to 15 experimental sessions, each session on a different day. Infants expressive language skills were assessed (REEL Scale 2, 1991) before submitting each of them to a learning task consisting of matching different animal-like figures in a matching-to-sample (visual-visual) conditional discrimination format. In a single-subject design, infants were taught four conditional discriminations: if A, then B; if A, then C; if D, then E; and if D, then F. The order of presentation and the left-right position of correct response were counterbalanced across training and testing trials. Once an infant learned these mixed relations under various reinforcers, the transitivity test was given. Equivalence was established when a child matched B and C, in as much as A had been the matching sample for both, and when a child matched E and F, both of which earlier had been paired to D. Every subject attained criterion on the four independent conditional discriminations and on the mixed training. 5 subjects who attained transitivity (at 80% or above), performed below chance level on at least one of the four symmetry tests. 8 out of our 10 subjects performed between 80 and 100 % correct responses in the transitivity tests. We found a significant negative correlation between the total number of trials to criterion during the conditional discrimination training and the combined receptive and expressive language quotient. Those infants with higher language-skill scores required fewer trials to complete the conditional-discrimination training. The results suggest that language skills play a role in stimulus equivalence formation.

There are at least three reasons why the concept of stimulus equivalence has captured the imagination of behavior analysts. First, the formation of equivalence classes is one of a range of cognitive phenomena, including those that denote concept formation, categorization, and rule-governed behavior that can be addressed and organized from a behavior-analytic perspective. Second, stimulus equivalence appears anomalous, unexpected, and emergent. And it is not immediately apparent that the emergence of stimulus equivalences are direct outcomes of operant learning. And, third, equivalence-class formation appears to be related to language development (see Devany, et al., 1986). In some way, the relations seen among stimuli in an equivalence class parallel the symbolic relations commonly said by cognitive developmental psychologists to be characteristic of language.

The Relation Between Thought and Language

The debate between Piaget and Vygotsky over the relationship between language and thought raised many questions about their developmental sequence in humans. Piaget deemphasized language by subordinating it to thought. He saw language primarily as a vehicle for expressing thoughts, not as a precursor to thought. On the other hand, Vygotsky (1962) argued that in the sensorimotor and early preoperational stages, thought and language develop independently. His position was that thought is prelinguistic and language is preintellectual. If educators knew the developmental point when symbolic representation and stimulus equivalence is possible, they would be able to teach more effectively. We assume that if infants are in fact capable of thought and of forming equivalence classes before language, we would not have to wait until their language has developed to start teaching them concepts, numbers, etc. Moreover, with definitive answers to how we develop thought and language and other human capabilities at an early age, we could structure more reliable tests as predictors of how infant conceptual progress would occur.

The Development of Language

Most normally developing infants can discriminate between the distinctive features of closely similar phonetic elements in speech, even at 1 month of age (Eimas, Siqueland, Einar, Jusczyk, & Vigorito, 1971). In natural circumstances, infants can recognize some symbolic word meanings as early as 6 to 8 months after birth (Bzoch & League, 1991). This recognition is followed shortly by demonstrated language skills for discriminating the meanings of most simple sentences by 9 to 12 months after birth. Receptive recognition vocabulary often rises to over 100 words by the first year of age, and the first intelligibly spoken words in expressive language normally tends to appear in the infant's repertoire between 9 and 14 months. Important for our research, however, is that expressive syntactic advances, typically, do not appear in the child's repertoires until after 18 months. For us, the key issue is to determine at which early developmental points stimulus equivalence-class formation can occur. The assumption that language skills can cause or facilitate successful equivalence-class formation needs further exploration.

Objectives of the Present Study

The main aim of this study is to explore how early in human development equivalence relations may appear and to examine the relation between stimulus equivalence and language competency. Previous studies have shown stimulus equivalence in normal children (Augustson et al., 1992), language-able mentally-retarded children (Devany et al., 1986, Saunders & Spradlin, 1993), and adults (Hayes, Thomas, & Hayes, 1989). An equivalence class is shown if the stimuli in the class show the three defining relations of reflexivity, symmetry, and transitivity (Sidman & Tailby 1982). Matching-to-sample procedures have been used to establish and test for these relations. In this test, the subject is presented with a sample visual stimulus, and given an array of three comparison visual stimuli underneath the sample stimulus. They are then asked to point to the comparison stimulus that goes with the one above, the sample. That is the child will learn that in the presence of the sample visual stimuli, if they pick a certain comparison visual stimuli they will be reinforced. Reflexivity is generalized identity matching, by matching a novel stimulus to itself under conditions of no reinforcement. That is to say, if a child is shown a picture of a car, the sample, and then asked which one of these comparisons go with the above, and they have to choose between a bat, tree or car. If the child picks the car, (comparison), to match the car, (sample), then this shows an identity relationship ("if A, then A"). Symmetry refers to the functional reversibility of the conditional relation; "if A, then B, and if B, then A". This means that if the sample visual stimuli is a car, and the comparisons are a house, truck and swimming pool, the child would be asked "Which one goes with the car?" If the child picks the truck then they will receive reinforcement for that behavior. This is called visual-visual conditional discrimination.

Some previous matching-to-sample tests have used auditory stimuli and asked the subject to match visual stimuli with it; this is called auditory-visual conditional discrimination. In a matching to sample test, in order to show transitivity at least 3 stimuli are required. If after the relations "if A, then B" and "if B, then C" have been taught with conditional discrimination trials, and then when tested for the relation "if A, then C" and the relation emerges without ever being explicitly trained with reinforcement or paired together, transitivity has been demonstrated.

Problems with Previous Studies with Young Children

From their findings, Devany, et al. concluded that language skills are related to the demonstration of stimulus equivalence. In their study, three groups of children were trained and tested on equivalence relations: (1) normally developing preschoolers, (2) children termed "retarded" showing expressive speech or signs, and (3) children termed "retarded" showing language-deficiency. Devany et al. reported the acquisition of visual-visual conditional discrimination and transitivity in every one of their language-able children (retarded and normal), whereas none of the language-deficient "retarded" children exhibited transitivity after having succeeded in the conditional discrimination training. Devany et al. concluded that the failure of the language-deficient children to form equivalence classes could not be explained on the basis of an inability of those children to learn conditional discriminations per se. That is because *all* the retarded/no language subjects in their study did learn the 4 conditional discriminations, while only those with language skill were able to show transitivity. Devany et al. attributed the inability of the retarded non-language children to form equivalence class to their language deficit. However, the Devany et al. (1986) study was based on the constraint sample of mentally-retarded infants that posits some potential confounds explain below. Based on the Devany et al. study, one can not arrive at a definitive conclusion about equivalence-class formation in 1- to 2-year-old children and its relationship to differential language (or cognitive) skills.

Our study was designed as a systematic replication of the Devany et al. (1986) and Auguston and Dougher (1992) studies. The same materials were used, mode of presentation, order of training the conditional relations and testing, as well as same reinforcement schedule. One divergence was that we used a more stringent criterion for training the conditional discriminations than did Devany et al. Instead of 9 correct out of ten responses, we used 9 consecutive correct responses to index the learning of each conditional discrimination. The reader is cautioned to look carefully when viewing the information given in the figures of the Devany et al. results. Their charts represent individual training and testing data for the 12 children (4 of whom were classified as "normal", 4 as "retarded-with-language," and 4 as "retarded-without-language"). Their classification, based on the Bayley Scale and the Stanford-Binet standardized intelligence test score, led to conclusions that: (1) the retarded/language skilled and the normal children all required fewer trials to complete the discrimination training than did the retarded children with no language skills; and (2) that all their language-able children performed better than the language-deficient children on the stimulus-equivalence test. However, when one examines the Devany et al. results showing the transitivity test scores of their 12 Ss, at least a question remains: Was transitivity indeed shown by the 8 language-skilled children? If we look at the results in the *first* block of every language-able subject on this test, we see that none of these subjects attained greater than 70 % correct responses in that first block, with many subjects scoring between the 50 and the 60 % level. One problem in interpreting the Devany et al. results is the way in which the data were calculated and plotted. In their graphs, each data point represented the number of correct responses out of the number of responses attempted, in a block of ten trials. This way of plotting presents an inconsistency in the representation of those data that can be misleading. For example, in a sample block of ten trials, there may have been only 4 trials with responses attempted and 6 trials with no attempts at all. Out of the 4 responses attempted, 3 may have been correct, resulting in a correct score of 75%. The enormous variability of the denominators yielding these percentages may provide a questionably-valid representation of performance during the transitivity tests. To correct for this problem, in the present study denominator values were held constant, always with 10 responses *attempted*. In general, subjects improved their performance from the first to the last transitivity test block of a

series. Even so, the answer as to whether or not infants subjected to this control procedure were able to form a new relation without implicit or explicit training remains unanswered.

Figure 1 shows results from our first subject (#1) when we tried to conduct a direct replication of the Devany et al. procedure. Notice that using their procedure, we were only able successfully to train the 2 conditional relations (A:B, D:E) independently and then mixed, in this 25-mos.-old language-skilled girl, with receptive and expressive language scores normal for her age. However, our subject was unable to demonstrate the emergence of the new relation during the transitivity test. We were about to run additional subjects to verify this contradictory finding, when Augustson's and Dougher's systematic replication of the Devany et al. study came to our attention (EAHB, 1992). By failing to replicate the Devany et al. results, Augustson and Dougher corroborated our concerns about the complexity of the Devany et al. procedure. But Augustson and Dougher left us with additional concerns about their own design. Their subjects were *never tested* for transitivity because they could not attain criterion during the mixed-training phase. We believe that this occurred because Auguston and Dougher used three comparison stimuli instead of two. Several reasons might explain why the subjects in these two studies failed to derive the new stimulus relation:

First, there is the possibility that children 24 mos. and younger may have difficulty in learning (and/or remembering) more than two conditional discriminations at a time before being tested for stimulus equivalence. In the Devany et al. design, all 4 conditional discrimination relations (A:B, D:E, A:C, D:F) were trained in a complex sequence, before the equivalence test was presented. Augustson and Dougher also trained the 2 conditional discrimination relations independently, before they introduced the mixed task. However, as in the aforementioned, their subjects were not able to reach criterion on the mixed task. Thus, Augustson and Dougher neither trained the next two relations nor tested for the transitive relation. Although Augustson and Dougher attempted to replicate the Devany et al. study, their results are not comparable to those of Devany et al. because those researchers made the learning more difficult for the children by adding a third comparison stimulus. They used an array of 3 comparison stimuli (one correct and two incorrect) on each trial. It is possible that this methodological change in the Auguston and Dougher design increased the complexity of the mixed training task beyond the skills of their 2 year-old subjects. Augustson and Dougher, in the second phase of their study, extended the mixed task and found that their subjects continued to perform at chance level even after two hundred trials, so they terminated the mixed training. But, even had the training been extended beyond 200 trials, one wonders if any of their child subjects would have succeeded in attaining criterion in the mixed training task, much less derived a new equivalence relation?

Is important to note that neither of those two studies tested for *symmetrical* responding. The possibility of such symmetrical responding was only inferred. In the present study we break down the complexity of the training sequence by training only one conditional-discrimination at a time and testing immediately for *symmetry* before training a new

relation. We expected that this addition in the design would permit us to assess for symmetry routinely after each subject has attained the response criterion in a conditional-discrimination training task. Thus, one purpose in our study was to test for symmetry, not to train it (see Figure 1).

Devany et al. used mentally-retarded subjects as a means of determining if language skills are prerequisite for stimulus equivalence. We believe it is difficult to determine if behavior patterns denoting "mental retardation" are associated with the failure of the subjects to derive equivalence relations (even when the groups are matched using mental age), or if the deficit in language skills *per se* could be responsible. Very importantly is that the data of Devany et al. revealed a significant difference in the number of unattempted trials between the retarded language and the retarded no-language groups. The retarded language group averaged 5 unattempted responses per child, whereas the retarded no-language group averaged 10. This could be another possibility for the failure of the retarded no language group to show the transitive relation. In addition, Devany et al. made only an informal assessment of the language and speech skills of their subjects. A formal method for assessing the language skills of infants was employed in our study.

All of these aforementioned issues prompted us to modify the procedure from that used in the earlier studies and to look for alternative tactics in the study of equivalence relations in very young children. Apart from concerns with methodological problems, our main long-term interest is to detect at which developmental points infants are able to demonstrate the emergent relations denoting symmetry and transitivity. We are interested in determining if any particular sequence in the conditional-discrimination training can facilitate the acquisition of equivalence relations. We are open to the possibility that stimulus-equivalence class formation would be manifested in young children at early developmental points, even before they show expressive language skill.

METHOD

Participants. 9 normal subjects, 7 males and 2 females aged 21 to 25 months, participated in this study. The mostly daily experimental sessions lasted between 20 and 30 minutes. The number of trials presented was not fixed, but usually involved at least 20 trials. The number of sessions ranged between 5 and 15.

Assessment Materials and Experimental Setting. Language Skills and Equivalence Relations. Before the first training session, each child's receptive and expressive language skills were assessed through the administration of The Bzoch-League Receptive-Expressive Emergent Language Scale (1991), for the Measurement of Language Skills in Infancy (REEL Scale 2). For reliability purposes, this language-skill assessment was made twice, one time by the experimenter alone, and a second time using the mother as an informant (with the experimenter reading each item to the mother). The average scores of the two informants were used for analysis of the data. This language test includes a 132-item checklist and uses observational information to identify the level of language skills in infants.

Two functional language systems, receptive and expres-

sive, are assessed directly. To the test developers, *receptive* language refers to the unified activity of all the sensory-neural associations and auditory-perceptual processes that are involved in the decoding and understanding of the intended meaning of oral languages (i.e., auditory comprehension). In contrast, *expressive* language refers to all of the underlying sensory-neural processes and also to the motor neural skills of the breathing, phonation, resonance, and articulation mechanisms of the body that are involved in communicating with others through the mediation of spoken symbolic languages.

Setting. In the experimental room, two video cameras recorded all activities. The subject sat at a table facing the experimenter. Six stimulus figures were used in the experiment. The tasks consisted of matching animal-like figures using a matching-to-sample format. Each stimulus figure was colored with one of six watercolor paints (red, brown, green, purple, yellow, and orange). Color assignment was random, except that all six colors had to be used in each stimuli set.

Procedure. A single-subject design was implemented. The children were taught four conditional discriminations: if A, then B; if A, then C; if D, then E; and if D, then F. This order was counterbalanced for half of the subjects. During training, either the A or D stimulus was presented as a sample with either B and E or C and F as comparisons. The left-right order of presentation of the comparison stimuli was counterbalanced across trials. Each child was trained and tested using a different stimulus set, made by randomly selecting from a pool of items. (The stimuli used in the equivalence test were identical to those used in the previous training phase, except that the sample stimuli were stimuli that previously had been comparisons during the conditional discrimination training.) Equivalence was indicated by matching B and C, inasmuch as A had been the matching sample for both, and by matching E and F, both of which earlier had been matched to D.

First, a child was taught to select B in the presence of A, the sample stimulus (A:B). The criterion for terminating training is nine consecutive (unprompted) correct responses. Once the child reached criterion on this relation, testing for the symmetrical relation (B:A) was conducted. Symmetry tests consisted of a block of ten trials in which responses were emitted. (Trials in which the child did not respond were not included in the ten trial blocks.)

After the symmetry test, the child was taught the second conditional discrimination, which was to select C in the presence of A (A:C). When the child reached mastery criterion on this task, testing for the symmetrical relation (C:A) was conducted. (Notice that at this point in the sequence the relation D:E was not trained as Devany et al. and Augustson and Dougher had; instead, A:C was trained). Then, a mixed task followed. In the mixed training the stimulus cards from both sets of the two previously acquired relations were mixed together and presented in a random order. Once the child reached criterion on this mixed task, the transitivity equivalence test was presented. The entire procedure was then repeated for the D:E, D:F relations.

At the start of each trial, the experimenter pointed to the sample stimulus and said, "See that one? Which one at the bottom goes with it?" Correct responses during training were reinforced with either praise, the blowing of bubbles, the

ringing of a bell, or the delivery of food (cereal, M&Ms). Incorrect responses were not reinforced. Physical prompting (guiding the child's hand to the correct choice) and visual prompting (placing the experimenter's finger on the correct choice) were used with some children at the beginning of training. Initially, a continuous reinforcement schedule was used and was gradually thinned until a programmed consequence was delivered only after every three or four correct responses. Reinforcement was not delivered for the target response during testing. Instead, the child was praised for cooperation, good sitting, and the like two or three times during a block of ten trials. The mixed training and the equivalence test were administered within the same session.

Recording. Behavior in each trial was scored as "correct", "incorrect" or "no response". A "correct" response is defined as touching the correct comparison stimulus while refraining from touching the incorrect comparison or the sample stimulus. An "incorrect" response is defined as touching the incorrect comparison or touching both the correct and incorrect comparisons. Any other behavior was scored as "no response."

RESULTS

As seen in Figure 1, the number of correct responses divided by ten (ten-trial block) was calculated to give a percentage of unprompted correct responses for each ten-trial block. The first block of each training period may reflect fewer than 10 trials with responses. In symmetry and transitivity testing periods, *all* blocks are comprised of 10 trials, each involving a response.

Figure 1 shows that all 9 subjects attained criterion (9 consecutive correct responses) on the four independent conditional discriminations and on the mixed training. Subjects required from 34 to 242 trials (mean=103) to learn the two relations (A:B, A:C,) and the mixed training. In contrast, the Devany et al. subjects required from 50 to 70 trials, with a mean of 68, to reach their response criterion. This difference in number of training trials to criterion may have to do with the fact that our subjects were less developmentally advanced, than the subjects of Devany et al. Our children ranged from 21 to 25 months and theirs from 25 to 52 months. The average number of correct responses for all subjects in the symmetry tests was 6.5. Surprisingly, 5 subjects who attained transitivity (at 80% or above), failed at least one of the four symmetry tests, that is, performed below chance level. Eight out of our 9 young subjects performed between 80 and 100 % correct responses in the transitivity tests.

A significant negative correlation of -.84 (at better than p < .01 alpha level) was found when the total number of trials to criterion during the conditional discrimination training was related *to* the combined receptive and expressive language quotient. That is, those children with higher language-skills scores required fewer trials to complete the conditional-discrimination training.

Figure 1.

PERCENT TRIALS CORRECT

BLOCKS OF TEN TRIALS
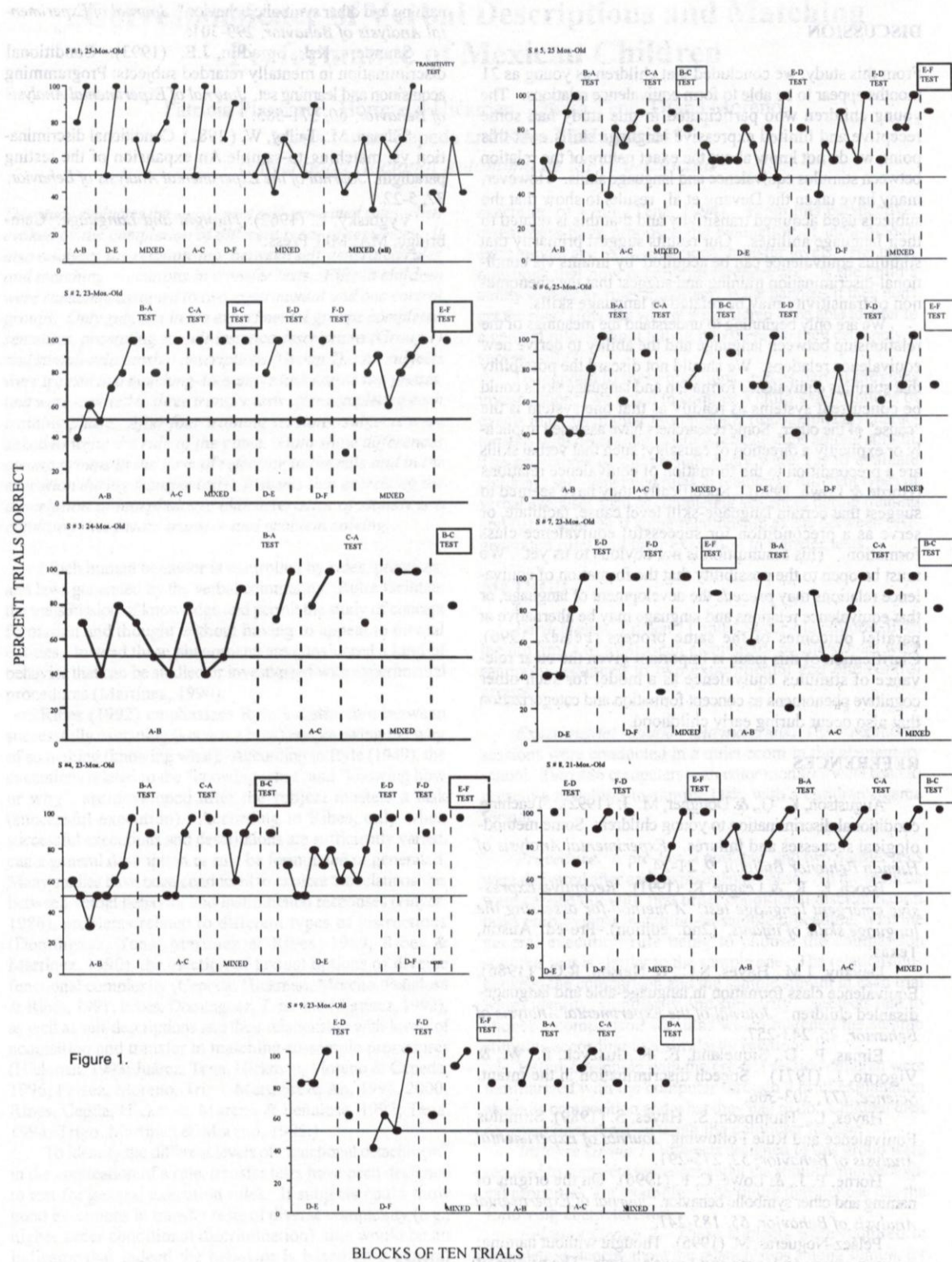
DISCUSSION

From this study, we concluded that children as young as 21 months appear to be able to form equivalence relations. The young children who participated in this study had some receptive and limited expressive language skills. At this point, we do not know about the exact nature of the relation between stimulus equivalence and language skills. However, many have taken the Devany et al. results to show that the subjects used acquired transitivity and that this is related to their language abilities. Our results suggest primarily that stimulus equivalence can be acquired by infants via conditional-discrimination training and suggest that the phenomenon of transitivity may be related to language skills.

We are only beginning to understand the meanings of the relationship between language and the ability to derive new equivalence relations. We should not discard the possibility that stimulus equivalence formation and language skills could be concurrent systems as readily as that one system is the "cause" of the other. Some researchers have assumed implicitly or explicitly a direction of causality, such that verbal skills are a precondition to the formation of equivalence relations (Horne & Lowe, 1996). Specifically, they have seemed to suggest that certain language-skill level cause, facilitate, or serve as a precondition for successful equivalence-class formation. This assumption is not obvious to us yet. We must be open to the possibility that the formation of equivalence relations may *precede* the development of language, or that equivalence relations and language may be alternative or parallel outcomes of the same process (Pelaez, 1996). Clarification of this issue is important given the clear relevance of stimulus equivalence as a model for such other cognitive phenomena as concept formation and categorization that also occur during early childhood.

REFERENCES

Augustson, K. G., & Dougher, M. J. (1992). Teaching conditional discrimination to young children: Some methodological successes and failures. *Experimental Analysis of Human Behavior Bulletin, 9*, 21-24.

Bzoch, K. R., & League, R. (1991). *Receptive-Expressive emergent language test: A method for assessing the language skills of infants.* (2nd. edition). Pro-ed: Austin, Texas.

Devany, J.M., Hayes, S.C., & Nelson, R.O. (1986). Equivalence class formation in language-able and language-disabled children. *Journal of the Experimental Analysis of Behavior, 46*, 243-257.

Eimas, P. D., Siqueland, E. R., Juszyck, P. W., & Vigorito, J. (1971). Speech discrimination in the infant. *Science, 171*, 303-306.

Hayes, L., Thompson, S., Hayes, S. (1989). Stimulus Equivalence and Rule Following. *Journal of Experimental Analysis of Behavior, 52*, 275-291.

Horne, P. J., & Lowe, C. F. (1996). On the origins of naming and other symbolic behavior. *Journal of Experimental Analysis of Behavior, 65, 185-241.*

Peláez-Nogueras, M. (1996). Thought without naming: A commentary on Horne and Lowe's article "The origins of naming and other symbolic behavior." *Journal of Experimental Analysis of Behavior,* 299-301.

Saunders, K.J., Spradlin, J.E. (1993). Conditional discrimination in mentally retarded subjects: Programming acquisition and learning set. *Journal of Experimental Analysis of Behavior, 60,* 571-585.

Sidman, M., Tailby, W. (1982). Conditional discrimination vs. matching-to-sample:An expansion of the testing paradigm. *Journal of the Experimental Analysis of Behavior, 37,* 5-22.

Vygotsky, L. (1962). *Thought and Language.* Cambridge, MA: MIT Press.