Sex Roles, Vol. 24, Nos. 3/4, 1991

Is There Coherence Among the Cognitive Components of Gender Acquisition?[1]

Barbara E. Hort,[2] Mary Driver Leinbach, and Beverly I. Fagot
University of Oregon

This study examined the relationships among five measures that assess various cognitive components of the child's acquisition of gender. At around 2 years of age, children were given a task assessing their ability to accurately label as "a boy" or "a girl" some head-and-shoulders pictures of boys and girls. At 4 years of age, these children were given tasks measuring (1) the degree to which they found gender a salient parameter of categorization, (2) the amount of gender-related knowledge they could display (SERLI-SRD), (3) the degree to which their preferences were gender-typed (SERLI-SRP) and (4) the accuracy of their memory for gender-typed information. There was no consistent pattern of relationship among the children's scores on these five tools for measuring gender acquisition. Our findings suggest that gender is a multidimensional construct in children's development, and thus these results challenge the undimensional manner in which gender is repeatedly addressed in developmental theory and research.

The most cursory review of the developmental literature will reveal the predilection of many researchers to refer to "sex role acquisition" as if it were a sin-

[1]This research was funded in part by a federal biomedical research grant (No. 2507RR07080) awarded to the first author from the Department of Psychology at the University of Oregon, and by Grant No. RO1HD175751 from the National Institute of Child Development awarded to the third author.

We wish to thank Margo Lane, Rachel Robertson, Jacqui Oucalt, Laura Baker, Dale Roberts, and the teachers, parents, and assistants of the Child Research Laboratory at the University of Oregon in Eugene, Oregon, for their invaluable contribution to this research.

[2]To whom reprint requests should be addressed at Department of Psychology, Reed College, Portland, Oregon 97202.

0360-0025/91/0200-0195$06.50/0

gle component of psychological development. This penchant for a univariate conceptualization of gender development is particularly troubling in light of the large body of evidence indicating that the child's acquisition of gender is most accurately compared to an intricate puzzle that the child pieces together in a rather idiosyncratic way. Specifically, the research on gender development increasingly shows that in order for the child to achieve a sensible model of gender, s/he must integrate many different components into a working conceptual and behavioral framework that revolves around the definitions and implications of "male" and "female." Gender-typed toy preferences, playmate preferences, aggression levels, and patterns of communication are among the numerous pieces in a child's gender puzzle, as are the cognitive components of gender labeling, gender stereotyping, and gender understanding. The wide variance among children's development of these components makes it difficult to understand the insouciance with which some researchers refer to sex role as a single developmental variable.

The recurring references to sex role development as a coherent developmental component are additionally problematic in the absence of empirical evidence to suggest that any significant intercorrelation exists among the various pieces of the gender puzzle. For example, Sears, Rau, and Alpert (1965) assessed the degrees of correlation among several behavioral and personality-based measures of gender acquisition. In their study, Sears et al. administered several tests of sex role development to a group of 40 children, 21 boys and 19 girls, of mean age 4.95 years. These tests included three personality measures—the "It" scale (Brown, 1956), several doll-play scales, and a observation rating that classified boys from "sissy" to "macho" and girls from "coquette" to "tomboy." The researchers also administered three behavioral measures to their subjects—a sex-typed play area usage scale, toy picture preference test (Fauls & Smith, 1956) and a toy play preference test. Sears et al. found (1) no significant pattern of correlation among the personality-based measures of gender development, (2) only a slight degree of correlation among the behavioral measures of gender development (whose minimal significance was carried by the data from the girl subjects), and (3) no significant pattern of corrleation among the personality- and behavior-based measures.

More recently, researchers studying sex role development have focused on childrens' performance on cognitive measures of gender acquisition. Once again, the research indicates that children must assemble a diverse number of component pieces—such as gender labels, gender stereotypes, accumulation and recall of gender-related knowledge, and gender-related evaluative processes—in order to construct a coherent notion of gender (cf. Huston, 1983). It also appears from this research that, in addition to being numerous, the cognitive gender puzzle pieces vary considerably in "shape"

from child to child. For example, while all children eventually organize gender-related information into two categories and then label those categories as some version of "male" and "female," it appears that the age at which the labels are assigned ranges between 21 and 44 months (Leinbach & Fagot, 1986). Moreover, while gender seems to hold great salience for children as a parameter for categorization (Martin & Halverson, 1981), Bem (1981) proposes that its salience is much greater for some children than it is for others. Similarly, the amount of information stored in the gender categories varies widely among children, as does the accuracy of their memory for this knowledge (cf. Huston, 1983). Finally, there appears to be a large variance among individual children's patterns of preference for gender-related material (e.g., Serbin & Sparfkin, 1986).

Although there has been no documented attempt to assess what, if any, degree of intercorrelation exists among the cognitive components of gender acquisition, some preliminary results of individual studies indicate the same disturbing (to proponents of sex role as a univariate concept) lack of coherence that Sears et al. found among the behavioral and personality-based measures. For example, Serbin and Sprafkin (1986) have shown that children who have stored more gender-related information do *not* necessarily demonstrate more gender-typed preferences. And while Martin and Halverson (1983) have found that children who have stored more gender-related information demonstrate better memory for gender-related information, this finding may be artifactual; how can children remember gender-related information that they never encoded in the first place? These individual studies are intriguing and yet their limited scope leaves us without a clear notion of how much the cognitive aspects of gender development are intercorrelated. Thus, the present research was designed to address the question, "How much coherence actually exists among the various cognitive components of gender acquisition?"

METHOD

Overview

The subjects in the present research were given measures on five components of gender acquisition. The first measure was a *gender labeling task* in which the subjects were tested on or near their monthly birthday, beginning at 18 months, in a procedure developed by Leinbach and Fagot (1986) to ascertain the age at which a child is able to assign the correct labels to head-and-shoulders pictures of girls and boys and men and women. The second measure, a *gender salience task* that was developed by Hort, Leinbach, and Fagot (1989) based on a task designed by Smith (1984), assessed

the degree of salience each child assigned to the categorical parameter of gender as opposed to the parameters of mood and activity. The third measure was the Sex Role Preference (SRP) factor included on the Sex Role Learning Inventory (SERLI; Edelbrock & Sugawara, 1978), which we called a *gender preference task* and used to measure the degree to which a child's gender-related preferences corresponded with standard gender stereotypes. The fourth measure was the Sex Role Discrimination (SRD) factor included on the SERLI (Edelbrock & Sugawara, 1978), which we called a *gender knowledge task* and used to measure the amount of traditional gender-related knowledge the child displayed. And the fifth measure, which we called a *gender memory task,* was an accuracy test of the child's memory for gender-related information, which we adapted from a measure used by Martin and Halverson (1983). The goal of this research was to compute the statistical relationships among the subjects' performances on these five measures, in order to test the notion of sex role as a univariate developmental entity as suggested, for example, by Bem's (1981, 1983) gender schema theory.

Subjects

The subjects of this study were drawn from a pool of 108 children who were regular participants in the longitudinal research conducted at the Children Research Laboratory in the University of Oregon in Eugene, Oregon, under the direction of the second and third authors. The children, nearly all of whom were white and middle class, were studied either in play groups and/or in individual task work over a period of four years. Due to the logistics of subject testing and the normal subject attrition over the four-year period of study, it was not possible to administer every task to every subject in the pool. Moreover, because the *gender salience task* and the *gender memory task* utilized the same set of line drawings, we were unable to administer both of these measures to any single subject without risking the subjects' confusion and artifactual performance. As a result, the *gender labeling task* was administered to 27 girls and 23 boys who were tested between the ages of 18 and 40 months, the *gender salience task* was administered to 27 girls and 35 boys at the age of 4 years, the *gender preference task* (SERLI-SRP) was administered to 26 girls and 30 boys at the age of 4 years, the *gender knowledge task* (SERLI−SRD) was administered to 26 girls and 30 boys at the age of 4 years, and the *gender memory task* was administered to 22 girls and 20 boys at the age of 4 years. Each statistical comparison of the children's performance on two of the five tasks was computed on the subset of children who received both the tasks being compared. All subjects were tested in a private room of the Child Research Laboratory by research assistants who were blind to the experimental hypotheses.

Gender Labeling Task

Materials. The stimulus materials for this task consisted of three parts, each of which contained 12 pairs of pictures held under plastic in a loose-leaf notebook. In the first part, each pair consisted of two familiar objects. In the second part, each pair included a stereotyped head-and-shoulders picture of a boy and another of a girl. In the third, each pair included a stereotyped head-and-shoulders picture of a man and another of a woman.

Procedure. All subjects receiving this task were first presented with the 12 pairs of familiar objects. For each pair, the subject was asked to point to the named object. Only subjects who performed this task successfully were allowed to continue. The subjects were then presented with the 12 pairs of boy-girl pictures. For each pair, the subject was asked in random order to point to either the boy or the girl. A score of 10 correct responses was beyond the level of chance responding and thus, was considered a passing score. This task was administered to each subject on his or her monthly birthday, beginning at the age of 18 months until s/he passed. The subjects were divided, according to the median split used by Fagot, Leinbach, and Hagan (1966), into dichotomous categories; *early labelers* were those children who passed the task on or before their 27 month birthday, and *late labelers* were those children who had not passed the task by their 27 month birthday.

Gender Salience Task

Materials. The stimulus materials for this task consisted of 12 sets of black and white cartoons of approximately the same size and shape on rectangular cards of equal dimension. Each of the 12 sets contained 9 cartoons, three of which were assigned to Experimenter 1, three of which were assigned to Experimenter 2, and three of which were assigned to the subject. The cartoons portrayed children whose features varied along three parameters: Gender (male and female), Mood (smiling and frowning), and Activity (running and sitting). Examples of these materials are shown in Fig. 1.

Procedure. The procedure used for this task was, like its stimulus materials, adapted from Smith's (1984) design. The subject was brought into the experimental room by Experimenters 1 and 2, and introduced to the categorization task with the following instructions:

> We're going to play the Matching Game today! I'm going to explain how we play this game. First, I get three cards and I lay them out in front of me, face up, like this. Then (Experimenter 2) gets three cards and s/he will lay them out face up, too. Then you get three cards and you lay them out in front of you, face up. All ready? OK.
>
> Now, I'm going to choose *this* card (Experimenter 1 chooses the card s/he holds which is identical to Experimenter 2's). Now, (Experimenter 2) gets to choose a card. (Experimenter 2 selects the card identical to the one chosen by Experimenter 1.)

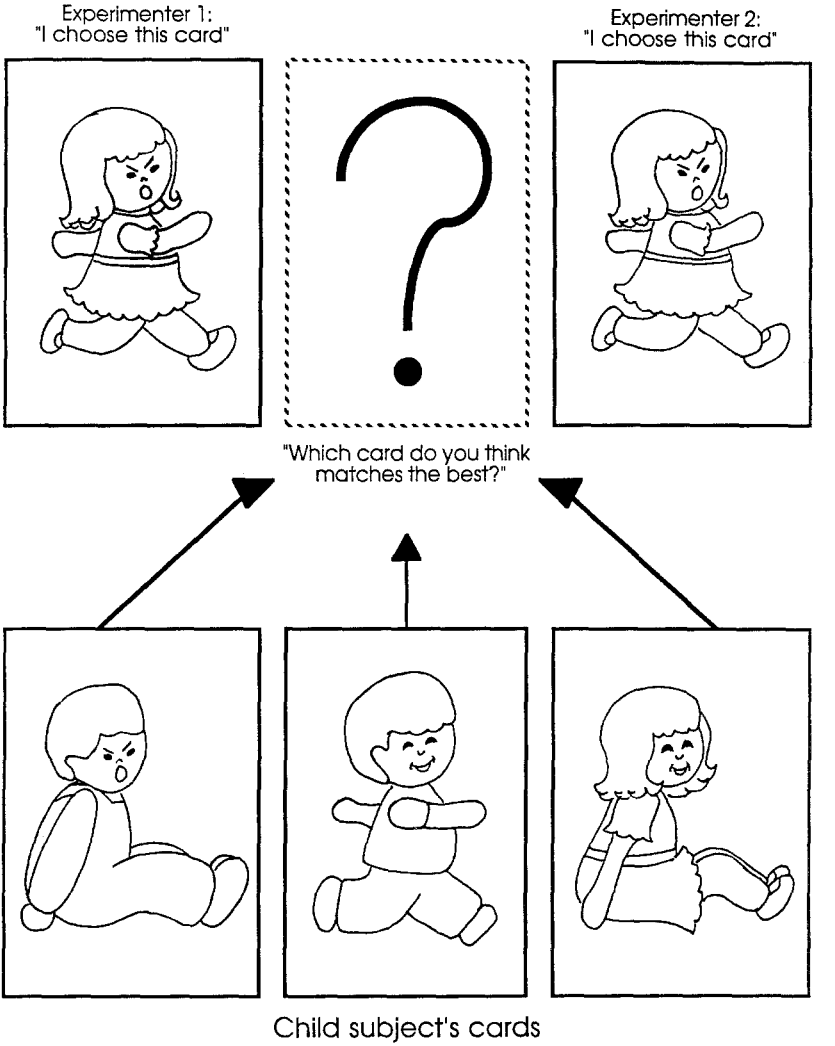

Fig. 1. Sample of materials used in the gender salience task.

> Now, you choose the card from your three cards which *you* think matches these two (identical) cards the best.

In this task, the subject could match the two identical cartoons on only *one* of the three parameters of Gender, Mood, or Activity. From the example above, in which the experimenters' cartoons portray a Frowning Girl Running, if the subject select the Frowning Boy Sitting, then the subject had selected a match based on Mood. If the subject chose the Smiling Boy Running, the match was based on Activity, and if s/he chose the Smiling Girl Sitting, the match was based on Gender.

This procedure was repeated for all 12 trials of the categorization task, with the order of the 12 trial sets counterbalanced across subjects. The basis of each choice made by the subject—Gender, Mood, or Activity—was recorded by Experimenter 1, while Experimenter 2 cleared the cards from the previous trials and arranged the cards for the new trial. At the end of the test session, each subject's score was computed as the total number of selections made on each of the three dimensions. Thus, each subject had three sums totaling to 12—one for matches on Gender, one for matches on Mood, and one for matches on Activity. Based on these scores, the subjects were divided into dichotomous categories—those subjects who chose matches based on Gender for 10 or more trials, and those subjects who based most of their matches on Mood or Activity or on none of the three parameters.

Once again, the cartoons in each participant's set of three were carefully arranged so that each of the two Experimenters' sets had one cartoon that matched the other experimenter's identically. The subject's set of three cartoons had *no* cartoons that matched any of the experimenters' exactly; instead, the subject's cartoons each matched the two identical experimenter cards on only *one* of the three parameters. For example, a set of nine cartoons in a trial of this study might be as follows (with salient parameters highlighted in bold face):

Experimenter 1	Experimenter 2	Subject
Smiling Boy Running	**Smiling Boy Running**	**Smiling** Girl Sitting
Frowning Boy Sitting	Frowning Boy Sitting	Frowning **Boy** Sitting
Smiling Girl Sitting	Smiling Girl Sitting	Frowning Girl **Running**

All 12 sets of nine cartoons were arranged in the above manner, with the permutations of gender, mood, and activity counterbalanced and equated for frequency of occurrence.

Gender Knowledge Task (SERLI-SRD)

Materials. The stimulus materials for this task were the two sets of line drawings developed for the SERLI (Edelbrock & Sugawara, 1978). The first set consisted of 12 line drawings of a child of the subjects' sex engaged in a gender-typed activity (6 male-typed and 6 female-typed), such as baking cookies and hammering a nail. The second set consisted of 12 line drawings of an adult of subjects' sex engaged in a gender-typed activity (6 male-typed and 6 female-typed), such as styling one's hair and working as a soldier.

Procedure. The SERLI-SRP task was administered and scored according to the system developed by Edelbrock and Sugawara (1978), who operationally define this test as measuring the degree to which the child's

classification of the various gender-typed objects agrees with sex role stereotypes of those items. The scores were adjusted to a scale ranging from 0 (*lowgender-stereotypic knowledge*) to 100 (*high gender-stereotypic knowledge*).

Gender Knowledge Task (SERLI-SRP)

Materials. The stimulus materials of this task was the set of 12 line drawings of gender-typed objects (6 male-typed and 6 female typed), such as a gun and a baby bottle, that were developed for the SERLI (Edelbrock & Sugawara, 1978).

Procedure. The SERLI-SRD task was also administered and scored according to the system developed by Edelbrock and Sugawara (1978), who operationally define this test as measuring the degree to which the child attempts to adhere to sex role stereotypes regarding appropriate masculine and feminine behavior. The scores were adjusted to a scale ranging from 0 (*low gender-stereotypic preferences*) to 100 (*high gender-stereotypic preferences*).

Gender Memory Task

Materials. The stimulus materials for this task were two sets of 12 line drawings. In each set, 6 were drawings of children (3 boys and 3 girls) engaged in same-gender activities (e.g., a boy going fishing), and 6 were drawings of children (3 boys and 3 girls) engaged in cross-gender activities (e.g., boy putting on eye makeup). The two sets of 12 drawings differed only in that the sex of the child performing each activity in the second set was the opposite of the child performing that activity in the first set. A sample of these materials is shown in Fig. 2.

Procedure. Although the line drawings used for the memory task were different from those used by Martin and Halverson (1983), the memory task was administered and scored in approximately the same manner. The only procedural modifications were that the children's memory was probed one hour after they had been shown the pictures, and the children's memory was probed for recollection of the 12 activities they had been shown and 4 neutral activities they had not seen. These changes were made to increase the number of activities and permutations shown to the children within the experimental time frame without taxing their attentional limits. Each child's performance score was adjusted to a scale ranging from 0 (*poor memory*) to 100 (*perfect memory*).

Fig. 2. Sample of materials used in the gender memory task.

RESULTS

The five tasks yielded five sets of scores to be compared. The scores from the *gender labeling task* produced a dichotomous variable whose categories were *early labelers* and *late labelers.* The scores from the *gender salience task* also produced a dichotomous variable whose categories were *gender salience* and *other/no salience.* The *gender preference task* (SERLI-SRP) produced a range of adjusted scores from 0 (*nongender stereotypic preferences*) to 100 (*gender-stereotypic preferences*). The *gender knowledge task* (SERLI-SRD) produced a range of adjusted scores from 0 (*low gender knowledge*) to 100 (*high gender knowledge*). And the *gender memory task* produced a range of adjusted scores from 0 (*poor memory*) to 100 (*perfect*

memory). Because some of these variables were dichotomous and others were continuous, the statistical test used for each planned comparison was dictated by the nature of the two variables being compared. These planned comparisons are summarized along with their statistical values in Table I.

The scores from the *gender labeling task* and the *gender salience task,* both of which were dichotomous, were compared using a Fisher's exact test and found not to be significantly related, indicating that children who find gender to be a salient means of categorization are not necessarily those who learned to label gender categories at an early age.

The scores from the *gender labeling task,* a dichotomous variable, were compared with the continuous scores of the *gender preference task,* the *gender knowledge task,* and the *gender memory task* by means of three standard *t* tests. Again, there were no findings of significant difference, indicating that children who learn to label the gender categories at an early age do not necessarily shown more traditionally gender-typed preferences, nor do they appear to have stored more gender-typed knowledge in those categories, nor do they demonstrate better memory for gender-typed information than children who learned to label the gender categories at a later age.

The continuous scores from the *gender preference task,* the *gender knowledge task,* and the *gender memory task* were compared using three Pearson tests of correlation, with no significant correlation among children who report gender-typed preferences, children who display a larger amount of gender-typed knowledge, and children who show good memory for gender-type material. It should be noted that the latter results are in contrast to Martin and Halverson's (1983) findings that children who possess more gender-typed knowledge display better memory for gender-related material. This disparity may be due to the fact that Martin and Halverson waited a week between the initial presentation of the materials and the probed memory recall, while the exigencies of our experimental design permitted a delay of only an hour. If so, it would imply that the correlation between gender-typed knowledge and memory for that knowledge varies among tasks, depending upon whether the memory requested is short-term or long-term.

Finally, two standard *t* tests were used to compare the dichotomous scores on the *gender salience task* with the continuous scores from the *gender preference task* and the *gender knowledge task.* There was no significant relation between the *gender salience task* scores and the *gender preference task* scores, indicating that children who find gender to be a salient parameter for categorization do not necessarily display gender-typed preferences, a finding contradictory to the premises of gender schema theory. However, there was a significant relation between the *gender salience task* scores and the *gender knowledge task* scores, indicating that children who find gender to

Table I. Planned Comparisons of the Five Cognitive Components of Gender Acquisition Tested in the Present Research

	Gender labeling (early vs. late)	Gender salience (gender vs. other)	Gender preference (0–100)	Gender knowledge (0–100)	Gender memory (0–100)
Labeling	—				
Salience	$p > .30$ (Fisher's)	—			
Preference	$p > .40$ [$t(24) = .83$]	$p > .90$ [$t(32) = .00$]	—		
Knowledge	$p > .15$ [$t(26) = 1.41$]	$p < .01$ [$t(32) = 2.81$]	$p > .35$ [$r(56) = -.10$]	—	
Memory	$p > .40$ [$t(13) = .91$]	na	$p > .35$ [$r(18) = .10$]	$p > .40$ [$r(20) = .08$]	—

be a salient parameter for categorization display more gender-typed knowledge than do children who find other categorical parameters more salient than gender.

GENERAL DISCUSSION

In the present research, we assessed the degree of relatedness among (1) the age at which a child could correctly label the gender categories, (2) the degree to which the child found gender a salient parameter for categorization, (3) the degree to which the child reported gender-stereotypic preferences, (4) the amount of gender-related knowledge the child displayed, and (5) the accuracy of the child's memory for gender-related information. In comparisons of all possible pairs of these factors (except the pair of gender salience and gender-related memory, which we were unable to compare due to overlapping task materials), we found evidence of only one significant correlation: children who found gender to be a salient means of categorization were also those children who displayed more gender-related knowledge. {The only other finding of note was the nonsignificant trend [$t(26) = 1.41$, $p > .15$] for early gender labeling to be associated with larger amounts of gender knowledge, a pattern that is consistent with findings by Fagot and Leinbach (1989) in which early gender labeling was correlated with greater gender knowledge (but not gender-typed preference) at age four.} There were no significant correlations among the age at which a children learns to label the gender categories, the degree of salience that gender holds for the child, how gender-typed the child's preferences are, how much s/he knows about gender, and how well the child remembers gender-related information.

The findings of this study definitely place in question the recurring propensity among researchers to refer to sex role acquisition as a single developmental component. In addition, these findings may have problematic implications for Sandra Bem's (1981, 1983) gender schema theory, which proposes (in terms of the gender puzzle metaphor) that the pieces of each individual's gender puzzle pieces fit together tightly and systematically. Our results suggest instead that children fit together the puzzle pieces of gender acquisition in a variety of loosely organized, idiosyncratic ways. Thus, our findings refute the conceptualization of "sex role" as a coherent univariate factor of cognitive development suggested in most developmental research and implied in Bem's gender schema theory. Rather, the results of our research support Huston's (1983) model of gender as a multivariate developmental construct that children acquire and employ in a variety of ways.

REFERENCES

Bem, S. L. (1983). Gender schema theory and its implications for child development: Raising gender-aschematic children in a gender schematic society. *Signs, 8,* 598-616.

Bem, S. L. (1981). Gender schema theory: A cognitive account of sex typing. *Psychological Review, 88,* 354-364.

Brown, D. G. (1956). Sex-role preference in young children. *Psychological Monographs, 70,* No. 421.

Edelbrock, C., & Sugawara, A. I. (1978). Acquisition of sex-typed preferences in preschool children. *Developmental Psychology, 14,* 614-623.

Fagot, B. I., & Leinbach, M. D. (1989). The young child's gender schema: Environmental input, internal organization. *Child Development, 60,* 663-672.

Fagot, B. I. Leinbach, M. D., & Hagan, R. (1986). Gender labeling and the adoption of sex-typed behaviors. *Developmental Psychology, 22,* 440-443.

Fauls, L. B., & Smith, W. D. (1956). Sex-role learning of five-year-olds. *Journal of Genetic Psychology, 89,* 105-117.

Hort, B. E., Leinbach, M. D., & Fagot, B. I. (1989). *A sex difference in the strategies for categorizing social information: Differential socialization or developmental decalage?* Unpublished manuscript, University of Oregon, Eugene.

Huston, A. C. (1983). Sex-typing. In P. H. Mussen (Ed.), *Handbook of child psychology* (4th ed., Vol. 4). New York: Wiley.

Leinbach, M. D., & Fagot, B. I. (1986). Acquisition of gender labels: A test for toddlers. *Sex Roles, 15,* 655-666.

Martin, C. L., & Halverson, C. F. (1981). A schematic processing model of sex typing and stereotyping in children. *Child Development, 52,* 1119-1134.

Martin, C. L., & Halverson, C. F. (1983). The effects of sex-typing schemas on young children's memory. *Child Development, 54,* 563-573.

Sears, R. R., Rau, L., & Alpert, R. (1965). *Identification and child rearing.* Stanford, CA: Stanford University Press.

Serbin, L. A., & Sprafkin, C. (1986). The salience of gender and the process of sex typing in three- to seven-year-old children. *Child Development, 57,* 1188-1199.

Smith, L. B. (1984). Young children's understanding of attributes and dimensions: A comparison of conceptual and linguistic measures. *Child Development, 55,* 363-380.