

Sex Stereotyping of Infants: A Review of Gender Labeling Studies¹

Marilyn Stern²

State University of New York at Albany

Katherine Hildebrandt Karraker

West Virginia University

Studies that have addressed the question of whether adults and children respond differently to male and female infants because of actual differences in the infants or because of preconceived sex stereotypes are reviewed, and the overall conclusions from these studies are evaluated. Twenty-three studies were identified in which neutrally clothed infant stimuli were labeled male in some conditions and female in other conditions. The strength and consistency of gender labeling effects on perceptions and behaviors in these studies were appraised. The results indicate that knowledge of an infant's gender is not a consistent determinant of adults' reactions, but more strongly influences young children's reactions. The implications of these findings for research on early sex role socialization are considered.

Most researchers agree that gender role socialization begins at the time of an infant's birth (e.g., Birns, 1976; Honig, 1983; Katz, 1979). Most parents are extremely interested in learning whether their newborn infant is a boy or a girl, and intentionally or not, this knowledge elicits in them a set of expectations consistent with consensual beliefs about sex role appropriate traits (Pogrebin, 1978; Rubin, Provenzano, & Luria, 1974). Evidence suggests that

¹An earlier version of this paper was presented at the meetings of the American Psychological Association, August 1986, Washington, DC. We would like to thank the anonymous reviewer whose comments facilitated improvements in this review.

²To whom reprint requests should be addressed at Department of Counseling Psychology, State University of New York at Albany, 1400 Washington Avenue, Albany, New York 12222.

these initial expectations, which form the basis of gender schemas (Bem, 1981), can have a powerful impact on parents' perceptions of and behavior toward infants (e.g., Fagot, 1978; Fitzgerald, 1977; Rubin et al., 1974; Smith & DGLISH, 1977). Gender contributes to the initial context within which adults respond to an infant, and thus may become an influential agent in the socializing process and development of the child's sense of self (Archer & Lloyd, 1985; Berndt & Heller, 1986).

Stereotyped expectations may influence gender role socialization and the acquisition of sex-typed behavior through a self-fulfilling prophecy process (Darley & Fazio, 1980; Deaux & Lewis, 1984). Preconceived gender-based expectations may cause the parent to elicit expected behavior from the infant and to reinforce expected behavior when it occurs, thereby confirming the parent's initial expectations. According to social learning theory, this process of early learning accounts for many of the observed sex differences in adult behavior (Maccoby & Jacklin, 1974). Over time, infants may be encouraged to emit expected sex-typed behaviors and eventually adopt sex role identities consistent with their parents' initial expectations.

Those concerned with the study of gender differences, gender role stereotyping, and gender role socialization have examined parent and infant behavior in an attempt to determine if these proposed processes occur. Differential parental attitudes toward and expectations regarding sons and daughters are evident immediately after birth. Parents in one study, for example, were asked to rate and describe their newborns shortly after birth when the primary source of information about the baby was his or her gender (Rubin et al., 1974). Although the infants did not differ on any objective measures, girls were rated as littler, softer, finer featured, and more inattentive than boys. Other studies have revealed that parents treat male and female infants differently. For example, boys are touched more frequently and handled more roughly before 3 months of age, whereas girls are touched more frequently after 6 months (Lewis & Weinraub, 1979). Girls also are verbally stimulated more than boys (Cherry & Lewis, 1976). Fagot (1978) observed that parents of toddlers reacted differently to boys' and girls' behaviors. Parents responded more positively to girls than boys when the toddlers played with dolls, and more critically to girls than boys when the toddlers engaged in large motor activity. Perhaps as a result of this differential parental treatment, male and female toddlers exhibit clear behavioral sex differences in their choice and manner of play. For example, toddlers' toy preferences match adult stereotypes (Fein, Johnson, Kosson, Stork, & Wasserman, 1975). Studies (Fagot, 1974; Smith & DGLISH, 1977) have found that girls play more with soft toys and dolls, dance significantly more, ask for help more, and dress up in adult clothes more than boys. In contrast, boys show more active play, show more play with transportation toys and blocks, and manipulate toys significantly more than girls.

These studies in the natural environment cannot determine whether existing sex differences in infant behavior elicit differential parental behavior or whether stereotyped parental expectations promote sex differences in infant behavior via a self-fulfilling prophecy process. A number of studies have attempted to resolve this issue using an experimental laboratory procedure that allows independent variation of an infant's actual sex and an adult's knowledge of the infant's sex. The experimental paradigm entails the use of one or more neutrally clothed infant stimuli that are labeled male for some subjects and labeled female for others. This approach holds infant behavior and appearance constant, and eliminates the possibility that inherent differences between male and female infants will be confounded with gender labeling effects. Subjects typically are given very little, if any, information about the infants other than gender. Since the actual infant stimuli remain constant or vary randomly across conditions, any differential responses to the infants as a function of gender label can be attributed to the influence of stereotypes rather than actual infant sex differences.

As a group, these studies suggest that adult and child responses coincide with culturally specified sex stereotypes associated with the gender label assigned to an infant and independent of actual infant gender differences. These studies have addressed how both perceptions and behaviors might be affected by expectations associated with the gender label assigned to an infant.

Many of these studies have been published in *Sex Roles* during the last decade. In addition, popular developmental textbooks and recent articles (e.g., Delk, Madden, Livingston, & Ryan, 1986; Honig, 1983; Paludi & Gullo, 1986; Shakin, Shakin, & Sternglanz, 1985) have cited these studies and their conclusions as evidence for early sex role stereotyping. Others (e.g., Herzog, Enright, Luria, & Rubin, 1982) have questioned whether gender label effects are as strong as some have suggested. Since a comprehensive review of this body of research has never been published, the purpose of the present effort was to review and reevaluate the results of this series of studies.

This review focuses first on the strength of the gender label effects reported not include gender labeling studies using child and adult stimuli. Biologically based sex differences are inconsistent (Birns, 1976) or nonexistent (Lewis & Weinraub, 1979) in early infancy, and yet parents treat boys and girls differently during this time period. This observation suggests that parent sex stereotypic expectations guide parent behavior. Consistent behavioral differences between boys and girls are evident in the preschool years (2–5 years) (Birns, 1976), suggesting that parental behavior may then be elicited by child behavior as well as, or instead of, by parental expectations. The study of the effects of knowledge of an infant's gender on adult reactions may help explain the genesis of sex differences readily observed in childhood and adulthood.

This view focuses first on the strength of the gender label effects reported in the literature. The overall strength of gender label stereotyping was assessed by examining the proportion of measures tested that showed significant gender label effects. We also considered whether gender label effects were stronger in some conditions than others. In particular, we assessed gender label effects when subjects either did or did not have an opportunity to interact with an infant, and whether gender labeling effects were stronger in parents, nonparent adults, or children, and in male or female subjects.

The second focus of the review is the consistency of findings, both across studies and in comparison with consensually acknowledged sex stereotypes. The consistency of findings across studies was evaluated by determining whether the same or similar measures showed effects in independent studies. The consistency of reported findings with widely accepted sex stereotypes was assessed by comparing the reported findings with other literature on the content of sex stereotypes.

METHOD

Selection of Studies

Twenty-three infant gender labeling studies were identified (see Table I). These studies were identified primarily through a computer search of the literature on infant gender label effects as well as by manually combing the contents of *Sex Roles* since its inception. Only those published studies were included in which infants were used as stimuli, infant gender label was manipulated, and subjects did not know the infants.

Characteristics of the Studies

The identified studies varied on several key dimensions. These dimensions included age, sex, and number of infant stimuli; method of infant stimulus presentation; characteristics of subjects; and type of response measures. Information on these dimension for each study is provided in Table I.

The infants used as stimuli ranged in age from 3 to 17 months. In 13 of the studies, subjects responded to only one infant, so the gender label effect was analyzed between subjects. In the remaining 10 studies, subjects responded to two or more infants, so the gender label effect was analyzed within subjects.

Infant stimuli were presented to subjects through videotapes (11 studies), photographs (2 studies), audiotapes (1 study), or live interaction (9 studies).

To simplify subsequent comparisons of studies, those studies involving no interaction between subjects and infants were combined. Thus, 14 studies were categorized as involving no interaction whereas 9 studies were categorized as involving interaction.

Children served as subjects in three of the studies, and adults served as subjects in the remaining studies. The adult subjects included exclusively parents in 7 studies (the pregnant women in the Bell and Carver study are considered parents), exclusively nonparents (typically college students) in 12 studies, and both parents and nonparents in the remaining study. The responses of children, parents, and nonparent adults are considered separately in some later analyses.

Three major categories of dependent variable were identified. The studies measured gender label effects on subjects' perceptions of the infant (assessed by several types of rating scales, most often bipolar adjective scales), behavioral responses (e.g., frequency and type of touching), and/or toy choice (i.e., preference for a masculine, feminine, or neutral toy).

All but three of the interaction studies (Sidorowicz & Lunney, 1980; Smith & Barclay, 1979; Zucker & Corter, 1980) used some type of behavioral assessment of the interaction between subjects and infants. Many of these studies also included a toy-choice measure. Rating scales were used exclusively in one interaction study (Smith & Barclay, 1979), and another two studies (Sidorowicz & Lunney, 1980; Zucker & Corter, 1980) reported the findings for toy choice as the only dependent variable. Two interaction studies (Bell & Carver, 1980; Seavey, Katz, & Zalk, 1975) used all three means of assessment.

All of the no-interaction studies used adjective rating scales to assess gender label effects. Condry, Condry, and Pogatshnik (1983) used reaction time to infant crying as well as rating scales to measure gender label effects.

Analysis

Since most of the studies identified for this review did not include information about effect sizes, it was not possible to employ standard meta-analysis techniques to determine the strength of gender stereotyping. Instead, gender label effects were compared across groups of studies by examining proportions of significant gender label main effects. Proportions were calculated by dividing the number of statistically significant gender label main effects by the total number of gender label main effect analyses conducted in a particular type of study or with a particular type of variable. These proportions are expressed below as percentages. When appropriate, proportions were compared by calculating a *Z* test of proportions. Although toy choice was sometimes analyzed using multiple measures in any one study, it is consi-

Sex Stereotyping of Infants

Meyer & Sobieszek (1972)	1	17	M	2	Videotape	19 M ^c 22 F ^c	25 M 19 F		0/24
Seavey et al. (1975)	1	3	F	1	Interaction		14M ^b 14 F ^b	0/10	0/8
Sidorowicz & Luney (1980)	2	3-11	F	1	Interaction		25 M 25 F		7/10
Smith & Barclay (1979)	1	4	M	1	Interaction		14 M 18 F		
Smith & Lloyd (1978)	2	6	M	1	Interaction	32 F		2/11	Yes
Sobieszek (1978)	1	17	M	2	Videotape		18 M 41 F		0/17
Stern & Hildebrandt (1984)	3	9	M	1	Videotape		114 M 111 F		0/17
(Study 1)	3	9	F						
Stern & Hilbrandt (1984)	1	9	M	1	Videotape	48 F			0/17
(Study 2)	1	9	F						
Stern & Karraker (1988)	2	9	M	4	Videotape	32 F			1/18
(1988)	2	9	F						
Stern & Karraker (1989)	2	9	M	4	Videotape		50 F		0/18
(1989)	2	9	F						
Will et al. (1976)	1	6	M	1	Interaction	11 F		0/5	Yes
Zucker & Corter (1980)	12	3-5	M	1	Interaction		25 F		No
(1980)	13	3-5	F						

^aIf univariate analyses were reported, each measure is counted here. If only multivariate analyses were reported, each multivariate cluster is counted as one variable.

^bOnly subjects exposed to infants labeled male or female are included here.

^cThis study is categorized as a nonparent study.

dered a single comparison for our purposes. For rating scale and behavioral measures, if univariate tests were conducted, all individual measures are included in the count in Table I. Otherwise, multivariate sets or cluster were counted as individual variables.

Interactions between gender label and subject sex were examined to determine if gender stereotyping was more likely in male or female subjects. Interactions between gender label and other factors were not considered, since none of these other factors were examined in more than one or two studies.

The consistency of gender labeling effects across studies was assessed for each variable found to be influenced by the gender label manipulation in an individual study. Similar variables were identified in other studies, and we report whether or not gender label effects also were found for these variables. Finally, we compared the reported findings in the reviewed studies with the content of sex stereotypes as applied to adults and children, using several reports from the literature as well as "commonsense beliefs" (Archer & Lloyd, 1985) about sex stereotypes.

RESULTS

Gender Label Effects Across All Studies and All Variables

As might be expected, the majority of studies (13 out of 23, or 56.52%) reported at least one main effect involving the gender label manipulation (see Table I). Of the remaining 10 studies, 3 (Bell & Carver, 1980; Condry et al., 1983; Meyer & Sobieszek, 1972) reported interactions between gender label and other factors. Thus, 69.56% (16/23) of the studies reviewed reported at least one gender label effect.

Calculation of the proportion of significant main effects considering all variables used and all comparisons made reveal that only 46 out of 252, or 18.25%, of the effects were significant (see Table II). Although this propor-

Table II. Proportions of Effects Significant

	Behaviors	Toy choice	Rating scales
All studies	9/50 (18.00%)	4/8 (50.00%)	33/194 (17.00%)
No interaction			26/164 (15.85%)
Interaction	9/50 (18.00%)	4/8 (50.00%)	7/30 (23.33%)
Children			21/30 (70.00%)
No interaction			14/20 (70.00%)
Interaction			7/10 (70.00%)
Adults			12/164 (7.32%)
Parents	9/39 (23.08%)		1/47 (2.13%)
Nonparent adults	0/11 (0%)		11/117 (9.40%)
No interaction			11/109 (10.09%)
Interaction			1/155 (1.82%)

tion is significantly higher than the 5% that would be expected by chance ($Z = 9.65, p < .05$), it is also lower than might be expected given the standard experimental labeling procedure used in the studies and the emphasis placed on the positive findings in the literature. This proportion is also likely to be an overestimation of the total proportion of infant gender label effects since it is based only on published studies. Many studies that do not find significant effects are not published (Light & Pillemer, 1984).

Gender Label Effects on Rating Scales

An overall total of 194 rating scale items were assessed across 17 studies. Thirty-three significant main effects were reported (17.00%, see Table II). In studies involving no interaction with an infant stimulus, 15.85% of the calculated main effects were significant, and in studies involving interaction, 23.33% of the effects were significant. These proportions were not statistically significantly different ($Z = 1.00, p > .05$).

Three of the studies included in these calculations used children as subjects (see Table I). In these studies, 70.00% of the calculated main effects were significant. For the studies using adults, only 7.32% of the main effects were significant. This latter proportion is not significantly higher than the 5% that would be expected by chance ($Z = 1.36, p > .05$). Comparing the proportions for children and adults revealed that significantly more gender label effects were obtained from children than from adults ($Z = 8.40, p < .001$). A comparison of the proportion of effects for parents (2.13%) vs. nonparent adult (9.40%) was not significant ($Z = 1.62, p > .05$).

One of the studies using children involved interaction and the other two did not. The proportion of findings (70.00%) was the same for both types of studies. Proportions also were recalculated for adults for each method of presentation (see Table II). The proportions for the two methods of presentation were not quite significantly different ($Z = 1.92, .05 < p < .06$).

The rating scales used and the gender label effects on these scales are listed in Table III. Consistency across the three studies using children as subjects was found for several variables. Labeled boys were rated as bigger, stronger, and louder (or noisier) in all three studies, and as faster, meaner, and harder in the two studies that assessed these variables. The findings that labeled boys were rated as more mad and smarter were inconsistent (i.e., the effect was found in one study but not another). Consistency across studies could not be assessed for the findings that labeled boys were rated as better eaters, less beautiful, more cheerful, and more tough, because these variables were utilized in only one study.

Consistency across the studies of adult raters was poor. Labeled boys were rated as more angry in two studies, but not in a third. Labeled boys

Table III. Rating Scale Measures and Effects^a

Study	Ratings	Expected effect ^b	Gender label effect	
Bell & Carver (1980)	12 bipolar adjective scales (not listed)	—		
Condry & Condry (1976)	Pleasure (in response to 4 stimuli)	F		
	Anger (in response to 4 stimuli)	M	LB higher (jack-in-the box)	
	Fear (in response to 4 stimuli)	F	LG higher (jack-in-the box)	
	Activity (includes quiet/loud, moving/still, fast/slow)	M	LB higher	
	Potency (includes aggressive/passive, little/big, strong/weak)	M	LB higher	
	Evaluation (includes good/bad, ugly/pretty, friendly/unfriendly)	—		
	Condry et al. (1983)	Fragility	F	
Cowan & Hoffman (1986)	Big/little	M/F	LB bigger	
	Mad/scared	M/F		
	Fast/slow	—	LB faster	
	Strong/weak	M/F	LB stronger	
	Nice/mean	F/M	LB meaner	
	Quiet/loud	F/M	LB louder	
	Smart/dumb	—		
	Soft/hard	F/M	LB harder	
	Delk et al. (1986)	No. of activities rated as masculine	—	LB higher
		No. of activities rated as feminine	—	LG higher
Haugh et al. (1980)	Big/little	M/F	LB bigger	
	Mad/scared	M/F	LB more mad	
	Fast/slow	—	LB faster	
	Strong/weak	M/F	LB stronger	
	Nice/mean	F/M	LB meaner	
	Quiet/loud	F/M	LB louder	
	Smart/dumb	—	LB smarter	
	Awake/Sleepy	—		
	Soft/hard	F/M	LB harder	
	Happy/Sad	—		
	Fun/not fun to play with	—		
	Good/bad	—		
	Haviland (1977)	Interest	—	
Enjoyment		F	LGs more	
Surprise		F		
Fear		F	LBs more	
Anger		M	LBs more	
Distress		F	LBs more	
Shame		—		
Contempt		—		
Hildebrandt & Fitzgerald (1979)	Cuteness	F	LBs more	
	Pleasure (in response to 2 stimuli)	F		
Holman & Williamson (1979)	Anger (in response to 2 stimuli)	M		
	Fear (in response to 2 stimuli)	F		

Table III. Continued

Study	Ratings	Expected effect ^b	Gender label effect
Meyer & Sobieszek (1972) ^c	Independent (M)	M	
	Sturdy (M)	M	
	Assertive (M)	M	
	Aggressive (M)	M	
	Impulsive (M)	M	
	Active (M)	—	
	Inquisitive (M)	—	
	Confident (M)	M	
	Alert (M)	—	
	Affectionate (F)	F	
	Passive (F)	F	
	Cuddly (F)	F	
	Coy (F)	F	
	Cooperative (F)	F	
	Delicate (F)	F	
	Receptive (F)	F	
	Shy (F)	F	
	Demanding	—	
	Selfish	—	
	Calm	M	
	Precocious	—	
	Friendly	—	
	Happy	—	
Possessive	—		
Seavey et al. (1975)	Passivity	F	
	Happiness	—	
	Friendliness	—	
	Precocity	—	
	Responsiveness	F	
	Physical attractiveness	F	
	Warmth	F	
	Comfort with baby	—	
Smith & Barclay (1979)	Big/little	M/F	LBs bigger
	Cuddly/not cuddly	F/M	
	Good eater/bad eater	—	LBs better eaters
	Not beautiful/beautiful	M/F	LBs less beautiful
	Noisy/quiet	M/F	LBs noisier
	Strong/weak	M/F	LBs stronger
	Unfriendly/friendly	—	
	Calm/nervous	M/F	
	Cheerful/fussy	—	LBs more cheerful
	Tough/gentle	M/F	LBs more tough
Sobieszek (1978)	Same as first 17 adjectives from Meyer & Sobieszek (1972)		
Stern & Hildebrandt (1984) (Study 1)	Same as Stern & Karraker (1988) minus confidence rating		
Stern & Hildebrandt (1984) (Study 2)	Same as Stern & Karraker (1988) minus confidence rating		

Table III. Continued

Study	Ratings	Expected effect ^b	Gender label effect
Stern & Karraker (1988)	Big/little	M/F	LBs bigger
	Large featured/fine featured	M/F	
	Very cute/not very cute	F/M	
	Friendly/shy	—	
	Happy/sad	—	
	Fun/not fun to play with	—	
	Smart/dumb	—	
	Attentive/inattentive	—	
	Loud/quiet	M/F	
	Pleasant/cranky-fussy	—	
	Awake/sleepy	—	
	Well behaved/not well behaved	F/M	
	Cuddly/not cuddly	F/M	
	Strong/weak	M/F	
	Assertive/passive	M/F	
	Fast/slow	—	
	Liking	—	
Confidence in rating	—		
Stern & Karraker (1989)	Same as Stern & Karraker (1988)		

^aM, male; F, female; LB, labeled boy; LG, labeled girl.

^bThese expected effects are based on the adult stereotyping literature (see text), and indicate which adjectives are more likely to be attributed to adult males and females. Where no expected effect appears, no information was available on whether the adjective would be more likely attributed to one sex or the other, or the adjective was equally likely to be attributed to each sex.

^cThe sex designations in parentheses are available in the description of the study.

were rated as more potent in one study and bigger (one component of the potency rating) in another, but not as more aggressive or strong (other components of the potency rating) in other studies that used these measures. All other labeling effects were found in one study, but not in others that used the same or a similar scale. However, only one clearly contradictory set of findings emerged: Labeled girls were rated as more fearful in the study by Condry and Condry (1976), labeled boys were rated as more fearful in the study by Haviland (1977), and no gender labeling effect on fear was found in the study by Holman and Williamson (1979).

To assess whether the findings listed in Table III were consistent with the content of sex stereotypes held by adults about other adults, lists of adjectives designated by adult subjects as characteristic of men and women (from Bem, 1974; Rosenkrantz, Vogel, Bee, Broverman, & Broverman, 1968; Spence, Helmreich, & Holahan, 1979; Spence, Helmreich, & Stapp, 1975; Williams & Bennett, 1975) were matched with the adjectives used in the in-

infant gender labeling studies. The sex most likely to be attributed each characteristic is listed as the "expected effect" for each variable in Table III.³

Of the 21 infant gender label effects found in the three child rater studies, 16 were consistent with the expected effects derived from the adult stereotyping literature. Five of the effects occurred for variables for which no expected effect could be determined from the adult literature. Only 3 additional effects would have been expected. Thus 16/19 (or 84.21%) of the expected effects were significant. This proportion is not significantly higher than the 70.00% previously calculated ($Z = 1.13, p > .05$).

Of the 12 infant gender label effects found with adult raters, 7 were consistent with predictions from the adult stereotyping literature and 3 were contradictory to those predictions. The remaining two effects (from Delk et al., 1986) could not be predicted directly from the adult stereotyping literature, but were consistent with widely held sex stereotypes. Of the 152 other tests conducted in the adult studies, gender labeling effects would have been expected at least 84 more times, based on the expected effects derived from the adult stereotyping literature. Thus 9/96 (or 9.38%) of the expected findings for adults were significant. This proportion is not significantly higher than the 7.32% previously calculated ($Z = .59, p > .05$).

Gender Label Effects on Behavioral Measures

Fifty behavioral measures were assessed in seven studies. Six of these studies involved interaction and one did not. The latter study (Condry et al., 1983) used reaction time to a crying infant as one of the dependent measures, whereas the other studies used various measures of stimulation, activity, touch, and verbal interaction. All but one of the interaction studies were conducted with parents. As seen in Table II, across all of these studies, 9 significant main effects were reported, yielding a proportion of 18.00%, which is significantly higher than a chance level of 5% ($Z = 4.22, p < .05$). Because of small numbers, the only comparison that could be made across these studies was of parents and nonparents. Although effects were found only with parents (23.08% vs. 0% for nonparents), the difference was not quite statistically significant ($Z = 1.76, .05 < p < .10$).

³In some cases, synonyms were used to match adjectives. In cases where pairs of adjectives were used in the infant studies, only one member of the pair had to match an adjective from the adult literature for an expected effect to be designated for the pair. No expected effect was listed when a trait was rated as non-sex-typed or a matching adjective could be found in the adult studies. A list of the sources for the matches for each infant scale item is available from the authors.

The behavioral measures used in each study and the gender label effects found are shown in Table IV. Consistency of findings across studies was difficult to assess because there was little overlap in the measures used in the different studies. The only finding consistent across two studies was that labeled boys were encouraged to be active in the studies by Frisch (1977) and Smith and Lloyd (1978). Other gender label effects were either assessed in only one study or were not confirmed in other studies using similar measures.

Comparison of gender label effects on behaviors with generally accepted sex stereotypes required more subjectivity than was the case with the rating data. The findings that labeled girls received more verbalization, interpersonal stimulation, and nurturance play, and that labeled boys received more encouragement of activity and more whole-body stimulation, seemed consistent with culturally defined sex stereotypes as well as observation of actual parent-infant interaction (e.g., Cherry & Lewis, 1976; Fagot, 1974, 1978). The labeled gender effects on adult gaze and neutral facial expression are not easily interpreted within the framework of culturally acknowledged sex stereotypes. No further attempt was made to determine the expected gender label effects on the remaining behavioral measures because of the paucity of previous data on which to base such judgments.

Gender Label Effects on Toy Choice

In each of the eight interaction studies employing toy choice as a measure, subjects were provided with a minimum of three toys to use while playing with an infant. The toys used are listed in Table V. Each of the toys available was categorized as masculine, feminine, or neutral in most of the studies. Although the sex appropriateness of the toys in the Seavey et al. (1975) and Sidorowicz and Lunney (1980) studies was not specified by the authors, we can safely assume that a football is stereotypically considered more appropriate for boys and a doll is considered more appropriate for girls. Significant gender label main effects on frequency or duration of toy use were found in four of the eight studies (50%). Due to the small number of comparisons, no additional analyses could be conducted to compare toy choice across different types of subjects. All studies involving toy choice were interaction studies with adults.

Results were somewhat inconsistent, but not contradictory, across studies. Dolls were selected more for labeled girls in three studies but not in four others studies. Labeled boys were more likely to be given a football in one study but not another, and a hammer in one study but not another. All the gender label effects on toy choice were consistent with stereotypic beliefs about the sex appropriateness of the toys, with the exception of the greater use of a rattle (considered a neutral toy) with labeled boys in the Smith

Table V. Toy Choice Measures and Effects

Study	Toy ^a	Gender label effect
Bell & Carver (1980)	Hammer (M)	
	Truck (M)	
	Teething ring (N)	
	Goose (N)	
	Duck (with ribbon) (F)	
Culp et al. (1983)	Doll (F)	
	Duck (N)	
	Doll (F)	
Frisch (1977)	Train (M)	
	Male role toy choice Female role toy choice (Toys used: beachball, blocks, bracelet, Bobo clown, doll, doll's bottle, fireman's hat, hairbrush, hand mirror, wall mirror, toy mop, Tykebike, puppets, pot, spoon)	More with labeled boy
Seavey et al. (1975)	Small rubber football	
	Raggedy Ann doll Flexible plastic ring	
Sidorowicz & Lunney (1980)	Small toy football	More with labeled boy
	Doll	More with labeled girl
Smith & Lloyd (1978)	Teething ring (N)	
	Squeaky hammer (M)	More with labeled boy
	Stuffed rabbit wearing trousers and bow tie (M)	
	Doll (F)	More with labeled girl
	Squeaky Bambi (F)	
	Squeaky pig (N)	
Will et al. (1976)	Plush ball (N)	
	Hourglass-shaped rattle (N)	More with labeled boy
	Fish (N)	
	Doll (F)	More with labeled girl
Zucker & Corter (1980)	Train (M)	
	Small red & white dump truck (M)	
	Batman doll (M)	
	Holly Hobbit doll (F)	
	Small set of pots & pans (F)	
	Squeaky mouse (N)	
	Rattle with movable face on one side and mirror on other (N)	

^aWhen specified by the author(s), toys are designated as masculine (M), feminine (F), or neutral (N).

and Lloyd (1978) study. In many cases, however, sex-stereotyped toys were not used differentially with labeled boys and girls. The lack of any labeled gender effects on toy use in some studies (particularly Zucker & Corter, 1980) may have been due to the age inappropriateness of the gender-associated toys.

Comparisons of Measures

Gender label effects for adults were significantly more likely with behavioral measures (18.00%) than with rating scale measures (7.32%) ($Z = 2.22, p < .05$). Effects on toy choice also were frequent (50%). These findings suggest that gender labeling affects adults' interactions with infants more than it affects their reported perceptions of infants.

Interactions Between Gender Label and Subject Sex

Seven interactions between gender label and subject sex were found in five studies. The content of these interactions was examined to determine if gender labeling effects were more likely in one sex or the other. Two of the interactions revealed a same-sex preference in subjects' responses. Haugh et al. (1980) reported that female children rated a labeled female infant as more fun to play with and that male children rated a labeled male infant as more fun to play with. Mayers and Sobieszek (1972) reported that adults attributed more positive and neutral qualities to an infant labeled their own sex.

Of the remaining five interactions, one indicated a gender label effect for men only whereas the other four indicated gender label effects for women only. Culp, Cook, and Housley (1983) reported that men, but not women, smiled more at a labeled boy than at a labeled girl. They also found that women, but not men, engaged in more verbal interaction with a labeled girl than with a labeled boy, and used a doll more with a labeled girl than with a labeled boy. Haviland (1977) found that women perceived more negative emotional expressions in labeled boys and more positive emotional expressions in labeled girls. These findings also suggest an occasional same-sex preference in subjects' responses.

DISCUSSION

This review of the infant gender labeling literature reveals that labeling an infant male or female does, in some cases, elicit sex-stereotypic responses from adults and children. The effects of this manipulation are not very strong, however. Overall, only about 18% of the variables assessed across a variety of studies were influenced by gender labeling. Although this number is greater than chance, it is well below a majority of the findings. In addition, many of the studies frequently cited as evidence for infant gender label effects (e.g., Meyer & Sobieszek, 1972; Seavey et al., 1975; Sobieszek, 1980) do not

report any significant gender label main effects. The results of the present review suggest that caution be exercised in assuming that infant gender label studies provide strong evidence for early gender role socialization.

Findings often were inconsistent across studies. Variables influenced by gender labeling in one study frequently were not influenced by the same manipulation in another study. Although these discrepancies could result from variations in power across studies due to sample size differences and variations in procedures, an examination of the sample sizes of the studies (see Table I) and the similarities among studies that produced different results indicates that these explanations are unlikely. These inconsistencies in findings suggest that additional research is needed to determine the conditions under which specific gender labeling effects occur.

Even though gender label effects were found only occasionally in the reviewed studies, identified effects typically were consistent with culturally acknowledged sex stereotypes. Authors of the studies generally did not specify which variables they expected would be influenced by gender labeling and which would not, so the present analyses were based on the hypothesis that all variables measured would be influenced. A more detailed comparison of the particular rating scale variables included in these studies with other information on sex stereotypes revealed that specific directional effects of gender labeling could only be expected for about half of the variables. However, recalculating the proportion of findings using just the variables for which effects were expected only marginally increased the proportion, since some of the gender label effects found were not predicted. In addition several contrast effects were discovered, where male and female infants were rated opposite to sex-stereotypic expectations. These findings emphasize the need for more information on the content of sex stereotypes (Deaux, 1984; Ruble & Ruble, 1982) and a more hypothesis-directed selection of variables in studies of infant gender labeling.

Additional analyses revealed that children responded much more strongly than adults to the gender labeling manipulation. Parents did not respond significantly differently from nonparents, and whether or not subjects interacted with an infant stimulus did not significantly influence responses. The children in the three studies reviewed were all between 2 and 6 years of age, a time when many children are forming their own gender identity. During this time children learn sex role stereotypes as cognitive-structural rules for understanding and organizing their social environment (e.g., Eaton, 1983; Reis & Wright, 1982). These children are therefore apt to perceive gender-related stimuli in more extreme and inflexible terms than adults do in order to maintain their self-identity (Eaton, 1983).

Although the findings on children's responses to gender-labeled infants are interesting, most infant socialization is controlled by adults rather than by older children. The studies of adults' reactions to gender-labeled infants

revealed both a generally low incidence of gender labeling effects and a tendency for behavioral measures (including toy choice) to be more strongly influenced by the gender label manipulation than were rating scale measures. Inconsistencies between behavioral and rating scale measures frequently have been identified in studies of sex stereotyping (e.g., Culp et al., 1983; Fatog, 1978; Will, Self, & Datan, 1976), as well as in studies of other social psychological processes (Ajzen & Fishbein, 1977; Wicker, 1969). Although these findings present a potential threat to the self-fulfilling prophecy model, the findings also suggest that the relationship between adult cognitions and behaviors relevant to infant gender role socialization may be more complex than previously has been conceptualized. For example, the particular sex-stereotyped cognitions that precede differential behavior toward male and female infants may not have been assessed in these studies.

The few gender label effects on adult ratings in the reviewed studies indicate that knowledge of an infant's sex is most likely to influence adults' interpretations of ambiguous infant behavior (Condry & Condry, 1976; Haviland, 1977), perceptions of infant physical characteristics (Condry et al., 1983; Hildebrandt & Fitzgerald, 1979; Stern & Karraker, 1988), and beliefs about appropriate infant activities (Delk, Madden, Livingston, & Ryan, 1986), but is unlikely to affect attributions of infant personality traits. The gender label effects on adult behaviors and toy choices suggest that adults are motivated to direct infant behavior toward sex-stereotyped activities. In combination, these findings suggest that knowledge of an infant's sex elicits a general stereotype that influences adult perceptions of the infant only when other information is absent or ambiguous (Darley & Gross, 1983), and most strongly affects the adult's conception of how the infant should act (now or in the future). This latter conception leads the adult to direct the infant toward sex-typed behavior.

Another consideration in interpreting this literature is that an infant's sex is only one influence on an adult's reaction to the infant. The adults in many of the reviewed studies may have reacted most strongly to the actual behavior and physical characteristics (e.g., size, physical attractiveness, motor maturity) of the infant stimulus to which they were exposed. The constraints of a laboratory situation also may have limited the strength of the gender labeling manipulation. The generally well-educated subjects in most of these studies may have studiously avoided reacting to infants in a sex-stereotypic fashion. In addition, only very brief exposure to infant stimuli are possible using this paradigm, and noticeable sex stereotyping effects may require more frequent or prolonged encounters.

Clearly, much work remains to be done to produce a complete understanding of how sex differences in children's behavior emerge. The present results suggest that knowledge of an infant's sex has subtle rather than pervasive effects on the cognitions and behaviors of adults. Future research must

become more process oriented (Deaux, 1984) in order to define more precisely how adults influence the emergence of early sex differences. The scope of gender labeling studies must be expanded to test hypothesized relations between gender knowledge and adult cognitions about and behavior toward infants. Moderating influences on gender label effects on perceptions and behaviors, including both infant and adult characteristics, also should be explored. Future research on early sex role socialization should consider modifications in the gender-labeling paradigm to make the assessments more reflective of natural interactive processes, to include measures known to differentiate between adult behavior toward boys and girls, and to consider other means of increasing the sensitivity of the procedure. We may also need to go beyond the gender labeling paradigm, given the inherent limitations of the technique and the limited knowledge gained from it so far. Researchers may need to recognize that active sex role socialization may not take place during brief laboratory encounters with unfamiliar infants, but only during recurring interactions with infants in the natural environment.

REFERENCES

- Ajzen, I., & Fishbein, M. Attitude-behavior relations: A theoretical analysis and review of empirical research. *Psychological Bulletin*, 1977, *84*, 888-918.
- Archer, J., & Lloyd, B. *Sex and gender*. Cambridge, London: Cambridge University Press, 1985.
- Bell, N., & Carver, W. A reevaluation of gender label effects: Expectant mothers' responses to infants. *Child Development*, 1980, *51*, 925-927.
- Bem, S. The measurement of psychological androgyny. *Journal of Consulting and Clinical Psychology*, 1974, *42*, 155-162.
- Bem, S. Gender schema theory: A cognitive account of sex typing. *Psychological Review*, 1981, *88*, 354-364.
- Berndt, T., & Heller, K. Gender stereotypes and social inferences: A developmental study. *Journal of Personality and Social Psychology*, 1986, *50*, 889-898.
- Birns, B. The emergence and socialization of sex differences in the earliest years. *Merrill-Palmer Quarterly*, 1976, *22*, 229-254.
- Cherry, L., & Lewis, M. Mothers and two-year-olds: A study of sex-differentiated aspects of verbal interaction. *Developmental Psychology*, 1976, *12*, 278-282.
- Condry, J., & Condry, S. Sex differences: A study of the eye of the beholder. *Child Development*, 1976, *47*, 812-819.
- Condry, S., Condry, J., & Pogatshnik, L. Sex differences: A study of the ear of the beholder. *Sex Roles*, 1983, *9*, 697-704.
- Cowan, G., & Hoffman, C. Gender stereotyping in young children: Evidence to support a concept-learning approach. *Sex Roles*, 1986, *14*, 211-224.
- Culp, E., Cook, A., & Housley, P. A comparison of observed and reported adult-infant interactions: Effects of perceived sex. *Sex Roles*, 1983, *9*, 475-479.
- Darley, J., & Fazio, R. Expectancy confirmation process arising in the social interaction sequence. *American Psychologist*, 1980, *35*, 867-881.
- Darley, J., & Gross, P. A hypothesis-confirming bias in labeling effects. *Journal of Personality and Social Psychology*, 1983, *44*, 20-33.
- Deaux, K. From individual differences to social categories: Analysis of a decade's research on gender. *American Psychologist*, 1984, *39*, 105-116.

- Deaux, K., & Lewis, L. Structure of gender stereotypes: Interrelationships among components of gender label. *Journal of Personality and Social Psychology*, 1984, 46, 991-1004.
- Delk, J., Madden, R., Livingston, M., & Ryan, T. Adult perceptions of the infant as a function of gender labeling and observer gender. *Sex Roles*, 1986, 15, 527-534.
- Eaton, W. Gender understanding and sex-role stereotyping in preschoolers: Implications for caregivers. *Child Care Quarterly*, 1983, 12, 28-35.
- Fagot, B. Sex differences in toddlers' behavior and parental reaction. *Developmental Psychology*, 1974, 10, 554-558.
- Fagot, B. The influence of sex of child on parental reactions to toddler children. *Child Development*, 1978, 49, 459-465.
- Fein, G., Johnson, D., Kosson, N., Stork, L., & Wasserman, L. Sex stereotypes and preferences in the toy choices of 20-month-old boys and girls. *Developmental Psychology*, 1975, 11, 525-528.
- Fitzgerald, H. Infants and caregivers: Sex differences as determinants of socialization. In E. Donelson (Ed.), *Women: A psychological perspective*. New York: John Wiley, 1977.
- Frisch, H. Sex stereotypes in adult-infant play. *Child Development*, 1977, 48, 1671-1675.
- Haugh, S., Hoffman, C., & Cowan, G. The eye of the very young beholder: Sex typing of infants by young children. *Children Development*, 1980, 51, 598-600.
- Haviland, J. Sex-related pragmatics in infants' nonverbal communication. *Journal of Communication*, 1977, 27, 80-84.
- Herzog, E., Enright, M., Luria, Z., & Rubin, J. Do gender labels yield sex differences in performance, or is label a fable? *Developmental Psychology*, 1982, 18, 424-430.
- Hildebrandt, K., & Fitzgerald, H. Adults' perceptions of infant sex and cuteness. *Sex Roles*, 1979, 5, 471-481.
- Holman, J., & Williamson, A. Sex labeling: Adult perceptions of child behavior. *Australian Journal of Early Childhood*, 1979, 4, 41.
- Honig, A. Sex role socialization in early childhood. *Young Children*, 1983, 38, 57-70.
- Katz, P. The development of female identity. In C. Koop & M. Kirkpatrick (Eds.), *Becoming female: Perspectives on development*. New York: Plenum Press, 1979.
- Lewis, M., & Weinraub, M. Origins of early sex-role development. *Sex Roles*, 1979, 5, 135-153.
- Light, R., & Pillemer, D. *Summing up: The science of reviewing research*. Cambridge: Harvard University Press, 1984.
- Maccoby, E., & Jacklin, C. *Psychology of sex differences*. Stanford, CA: Stanford University Press, 1974.
- Meyer, J., & Sobieszek, B. Effects of a child's sex on adult interpretations of its behavior. *Developmental Psychology*, 1972, 6, 42-48.
- Paludi, M., & Gullo, D. The effect of sex labels on adults' knowledge of infant development. *Sex Roles*, 1986, 16, 19-30.
- Pogrebin, L. Non-sexist parenting at home and at school. In B. Sprung (Ed.), *Perspectives on non-sexist early childhood education*. New York: Teachers College Press, 1978.
- Reis, H., & Wright, S. Knowledge of sex-role stereotypes in children aged 3 to 5. *Sex Roles*, 1982, 8, 1049-1056.
- Rosenkrantz, P., Vogel, S., Bee, H., Broverman, I., & Broverman, D. Sex role stereotypes and self-concepts in college students. *Journal of Consulting and Clinical Psychology*, 1986, 32, 287-295.
- Rubin, J., Provenzano, R., & Luria, Z. The eye of the beholder: Parents' views on sex of newborns. *American Journal of Orthopsychiatry*, 1974, 44, 512-519.
- Ruble, D. N., & Ruble, T. L. Sex stereotypes. In A. G. Miller (Ed.), *In the eye of the beholder: Contemporary issues in stereotyping* (pp. 188-252). New York: Praeger, 1982.
- Seavey, C., Katz, P., & Zalk, S. Baby X: The effect of gender labels on adult responses to infants. *Sex Roles*, 1975, 1, 103-109.
- Shakin, M., Shakin, D., & Sternglanz, S. Infant clothing: Sex labeling for strangers. *Sex Roles*, 1985, 12, 955-964.
- Sidorowicz, L., & Lunney, G. Baby X revised. *Sex Roles*, 1980, 6, 67-73.
- Smith, C., & Lloyd, B. Maternal behavior and perceived sex of infant revisited. *Child Development*, 1978, 49, 1263-1265.

- Smith, K., & Barclay, L. Kindergarten children's descriptive sex-stereotyping of infants. *American Journal of Orthopsychiatry*, 1979, *49*, 514-517.
- Smith, P., & Daghli, L. Sex differences in parent and infant behavior in the home. *Child Development*, 1977, *48*, 1250-1254.
- Sobieszek, B. Adult interpretations of child behavior. *Sex Roles*, 1978, *4*, 579-588.
- Spence, J. T., Helmreich, R. L., & Holahan, C. K. Negative and positive components of psychological masculinity and femininity and their relationships to self-reports of neurotic and acting out behaviors. *Journal of Personality and Social Psychology*, 1979, *37*, 1673-1682.
- Spence, J., Helmreich, R., & Stapp, J. Ratings of self and peers on sex role attributes and their relation to self-esteem and conceptions of masculinity and femininity. *Journal of Personality and Social Psychology*, 1975, *32*, 29-39.
- Stern, M., & Hildebrandt, K. Prematurity stereotype: Effects of labeling on adults' perceptions of infants. *Developmental Psychology*, 1984, *20*, 360-362.
- Stern, M., & Karraker, K. Prematurity stereotyping among mothers of premature infants. *Journal of Pediatric Psychology*, 1988, *13*, 253-262.
- Stern, M., & Karraker, K. Modifying the prematurity stereotype: Impact of providing varied information. *Journal of Clinical and Social Psychology*, 1989, *8*, 1-13.
- Wicker, A. W. Attitudes versus actions: The relationship of verbal and overt behavioral responses to attitude objects. *Journal of Social Issues*, 1969, *25*(4), 41-78.
- Will, J., Self, P., & Datan, N. Maternal behavior and perceived sex of infant. *American Journal of Orthopsychiatry*, 1967, *46*, 135-139.
- Williams, J. E., & Bennett, S. M. The definition of sex stereotypes via the adjective check list. *Sex Roles*, 1975, *1*, 327-337.
- Zucker, K., & Corter, C. Sex stereotyping in adult-infant interaction: Some negative evidence. *American Journal of Orthopsychiatry*, 1980, *50*, 160-164.