

Differing Levels of Gender Salience in Preschool Classrooms: Effects on Children's Gender Attitudes and Intergroup Bias

Lacey J. Hilliard and Lynn S. Liben

The Pennsylvania State University

Developmental intergroup theory posits that when environments make social-group membership salient, children will be particularly likely to apply categorization processes to social groups, thereby increasing stereotypes and prejudices. To test the predicted impact of environmental gender salience, 3- to 5-year-old children ($N = 57$) completed gender attitude, intergroup bias, and personal preference measures at the beginning and end of a 2-week period during which teachers either did or did not make gender salient. Observations of peer play were also made at both times. After 2 weeks, children in the high- (but not low-) salience condition showed significantly increased gender stereotypes, less positive ratings of other-sex peers, and decreased play with other-sex peers. Children's own activity and occupational preferences, however, remained unaffected.

People vary along many dimensions such as age, race, political party, hair color, and favorite activities. Why do children use some dimensions rather than others to categorize people into social groups? And what (if any) consequences result from sorting people into those groups? These questions are at the center of a recent formulation of developmental intergroup theory (Bigler & Liben, 2006, 2007) that builds on several other intergroup approaches (Aboud, 2003; Cameron, Alvarez, Ruble, & Fuligni, 2001; Killen & McKown, 2005; Nesdale, 2004; Rutland, 1999; Rutland, Cameron, Milne, & McGeorge, 2005).

Developmental intergroup theory views the emergence of social stereotypes in children through constructivist lenses. Thus, as in other constructivist views of gender development (e.g., gender schema theory; see Martin & Halverson, 1981), children are

thought to be active agents in processing social stimuli. Much early research in the constructivist tradition addressed how gender schemata influence the way that children process material differently depending on whether the material is gender traditional versus nontraditional and on whether it is culturally linked to the child's own versus the other sex. Illustrative of the former is research showing that children tend to have difficulty in remembering nontraditional material either by distorting it so that it becomes consistent with gender stereotypes or by forgetting it entirely (e.g., Koblinsky, Cruse, & Sugawara, 1978; Liben & Signorella, 1980). Illustrative of the latter is research showing that children are more likely to attend to, remember, and enact material associated with their own sex (e.g., Grusec & Brinker, 1972; Martin, Eisenbud, & Rose, 1995; Signorella, Bigler, & Liben, 1997). In developmental intergroup theory, these constructive processes are examined for their role in forming and maintaining social-group categories that in turn become the basis for developing stereotypes and prejudices.

Bigler and Liben (2006, 2007) argued that the development and application of children's stereotyping and prejudice are influenced by the salience of groups within children's surroundings. They identified various factors that render groups more or less salient. One factor concerns the degree to

The data reported here were collected as part of the first author's master's thesis, conducted under the supervision of the second author. Portions of this work were presented in April 2008 at the Gender Development conference, San Francisco. We express our appreciation to the child-care directors, Janie Hilliard and Barbara Crosby, for their willingness to accommodate this project, and to the participating staff, teachers, children, and parents for their friendly cooperation. We are also grateful to the members of the Cognitive and Social Development Lab at Penn State for their many contributions to data coding and entry, to Rebecca Bigler for her insights during various phases of this research, and to the Faris endowment of Penn State's Department of Psychology for financial support.

Correspondence concerning this article should be addressed to Lacey J. Hilliard or Lynn S. Liben, Department of Psychology, Moore Building, The Pennsylvania State University, University Park, PA 16802. Electronic mail may be sent to hilliard@psu.edu or liben@psu.edu.

which group membership is visually or perceptually observable (e.g., skin color vs. political affiliation). Irrespective of its initial visibility, a particular quality may be made more visible by an environment that marks group membership in some way. For example, men and women may wear different clothing and hair styles that exaggerate their physical differences, members of different religious groups may wear different head coverings or religious symbols, and members of different political parties may don distinctive lapel buttons. In addition, groups may be made more salient through the use of different language and labels to refer to group members (Bigler, Jones, & Lobliner, 1997).

Also contributing to the impact of group divisions is the child's own drive to categorize: As a part of their constructive, self-motivated drive to understand their interactions with their social and physical environment, children attempt to categorize their early environment and then make sense of these categorizations (Waxman & Gelman, 1986; Waxman & Markow, 1998). Thus, when children recognize differences between groups and are made aware of those groups through implicit and explicit categorization and labeling, they are likely to use those bases to divide people into groups. They are then internally and externally driven to attach meaning to those categories. For example, if religion is made salient in a child's environment (e.g., via differential clothing and labeling), young children would be expected to be more vigilant than they otherwise might be in seeking to find what is associated with religion. Children may also treat categories as innate due to their tendency toward essentialism (Gelman, 2003). Thus, if the child's environment divides people by race, children would be expected to be more likely to infer that differences between races rest on fundamental differences between the groups.

Children also have a propensity toward an in-group and out-group bias (Abrams, Rutland, & Cameron, 2003; Killen, Margie, & Sinno, 2006), and thus once children have enacted their categorization processes, they tend to view their own group as superior and the "other" group as inferior. Intergroup approaches have been used to study and understand stereotypes and prejudices in children. Early work showed that once children are assigned to groups, they make positive attributions to their own in-group and negative attributions toward the out-group (Sherif, Harvey, White, Hood, & Sherif, 1961). Consistent with the core tenets of developmental intergroup theory, these intergroup processes are heightened when there is a physical

difference between groups (e.g., race; see Aboud, 1988). The importance of the distinction between in- and out-group attitudes has received recent attention, specifically in a call for measures that do not assume a reciprocal association between the two (Brewer, 1999). Recent investigators have also questioned the yoked nature of in-group preference and out-group derogation and have suggested that the two constructs are not inextricably linked (Aboud, 2003; Cameron et al., 2001; Kowalski, 2003).

One important arena in which the processes identified in developmental intergroup theory may operate is gender. Indeed, over a quarter century ago, Bem (1983) offered an analysis that is consistent with the ideas presented in developmental intergroup theory in that she argued that society's emphasis on gender categorization led to a host of dysfunctional and unnecessary sex differences. Despite this critique, children continue to encounter experiences that increase the salience of gender as when teachers greet their pupils with "Good morning, boys and girls" or when schools require that boys and girls wear different-colored robes at commencement ceremonies.

There has been only a small body of empirical research bearing on the impact of different levels of gender salience. Studies have shown the effect of gender labels on attributing traits to individuals (see Stern & Karraker, 1989), but few experiments have evaluated the role of using gendered language in the environment or have studied whether in-group preferences and out-group biases emerge as predicted by intergroup theories. In one of the few relevant studies, Bigler (1995) found that when teachers used gender to organize their elementary school classes, children showed greater gender stereotyping than did children in classrooms that avoided the functional use of gender. In this study, younger children with less advanced classification skills were more susceptible to the effects of the manipulation than older children (i.e., children with less advanced classification skills developed more negative attitudes than peers with more advanced classification skills).

Although not applied specifically to gender, Patterson and Bigler (2006) have provided evidence that even younger children are also susceptible to developing in-group biases based on group divisions. Specifically, they used a novel group paradigm in which preschool children were randomly assigned to different-colored shirts. They found that when the teacher made functional use of color groups (e.g., lining children up for recess by shirt

color), children developed significantly stronger in-group biases. However, it remains unknown if preschool children would be affected by gender labeling and organization in classroom environments. Research to address this issue is important not only because of the practical implications for preschool classroom structure but also because it concerns a category division—gender—that is already emphasized heavily in young children's lives. It is possible that in this larger societal context, preschool classroom manipulations would have little impact.

The evidence just discussed suggests that young children's gender attitudes are susceptible to environmental manipulations that make social-group categories highly salient. Research in the intergroup literature has largely focused on the consequences of sorting people into groups and how one's sense of identity toward a group influences intergroup relations (Tajfel & Turner, 1986). Investigators have shown that gender is prevalent as an in-group category and have reported children's gendered play patterns and own-sex preferences (Maccoby & Jacklin, 1987; Yee & Brown, 1992, 1994). Developmental intergroup theory draws on the fields of intergroup relations but provides a model to test *how* stereotypes are formed and exaggerated, specifically by manipulating the salience of a social-group category in children's environments. The current study expands on previous work of intergroup processes and novel group paradigms by incorporating a wide range of measures to assess children's intergroup preferences and peer play behavior in the domain of gender. More specifically, the purpose of the present study was to test whether increasing the salience of gender in preschool classrooms would lead children to develop more highly gender-stereotyped attitudes and whether there would be evidence of both increased in-group and out-group bias (i.e., favoring one's own gender group and avoiding the other-gender group).

To address these questions, children were studied over a 2-week period under one of two gender-salience conditions. In the low-salience condition, children continued to experience a preschool environment in which the teacher avoided making gender explicit. In the high-salience condition, children experienced a preschool environment in which the teacher highlighted gender by using gender-specific language (labeling groups and individuals with gender terms) and by using gender-based organization in the classroom (e.g., assigning boys and girls to separate bulletin boards, lining up boys and girls separately for recess). Based on developmental intergroup theory, we predicted that children in the

low-salience condition would show no discernible changes in their gender stereotypes over the 2-week period of the study, whereas children in the high-salience group would show a significant increase in their gender stereotypes (i.e., a decrease in gender flexibility).

If the increased salience of gender in the classroom was indeed found to strengthen children's gender stereotypes, developmental intergroup theory would predict that the high-salience condition should also yield stronger in-group and out-group bias. To test this latter prediction, we assessed children's interest in playing with peers of their own sex (in-group) and with peers of the other sex (out-group). Specifically, both prior to and following the classroom intervention, children were asked to rate their interest in playing with each of the other children in their class, and their average ratings for own- versus other-sex peers were compared across the two times. As an additional posttest measure, children were asked to select which of three groups of unfamiliar peers they would like to play with—one with only boys, one with only girls, or one with both boys and girls. Of interest was whether children in the high-salience classrooms would express greater reluctance to play with other-sex children than would children in the low-salience classrooms. As a third way to evaluate children's interest in own- versus other-sex peers, we collected observational data on children's peer play. We predicted that children in the high-salience (but not low-salience) classrooms would show a significant decrease in play with other-sex peers over the 2-week period. Measures in the present study were designed so that in- and out-group biases could be assessed independently (i.e., children could show positive in-group attitudes without showing negative out-group attitudes, and vice versa). Researchers have suggested a primacy of in-group preference compared to out-group derogation (Aboud & Amato, 2001; Brewer, 1999). If children's preference for in-group members is strongly established by the early preschool years whereas negative attitudes toward out-group members develop somewhat later, then an experimental manipulation in the preschool years would be more likely to increase negative attitudes toward out-group members rather than influence already-established positive attitudes to in-group members.

The predictions described earlier concern children's gender attitudes about others and the consequences of those attitudes (i.e., in-group and out-group biases). Another aspect of gender development concerns the gendered nature of children's

own interests, for example, whether children favor toys and educational or occupational pathways that are culturally viewed as masculine or feminine. Past research has suggested that children's gendered personal interests are generally resistant to short experimental interventions, even when those interventions have a significant impact on their beliefs about what others should do (e.g., Bigler, 1995; Bigler & Liben, 1990). To examine whether a similar pattern of results would hold here, we also collected pre- and posttest data on personal interests in culturally stereotyped activities and occupations.

Method

Participants and School Context

Participants were 57 children ranging in age from 3 years 1 month to 5 years 6 months ($M = 4.7$ years; $SD = 0.65$ years). Children were roughly evenly divided between two preschools, each enrolling approximately equal numbers of boys and girls (14 boys and 16 girls in one school; 14 boys and 13 girls in the other). The majority of participants were European American from middle-class backgrounds. Informed consent was obtained from the directors of each school and teachers in each classroom. All parents and children were informed about data collection (interviews and observations) by letter and given the opportunity to opt out of the research activities if desired. None of the parents or children declined to participate.

The two schools (four classrooms) were located in a mid-sized city in the Southwest and were similar along important dimensions (e.g., teacher education, class size, teacher-child ratio, populations served, policies implemented, and classroom arrangement). Despite similarities, however, there are always differences between any two schools and between any two classrooms within a given school and grade (e.g., different teachers, different peers). Thus, the design included both pre- and posttests on critical measures so that analyses could take into account any initial differences between children in the two conditions.

Classroom Conditions

Classrooms in the high-salience condition were designed to make gender salient by exposing children to frequent uses of gender classifications in their environment. Following procedures used by Bigler (1995), teachers were encouraged to make

frequent use of gender through physical separation (e.g., lining children up by sex), classroom organization (e.g., posting separate boys' and girls' bulletin boards), and gender-specific language (e.g., "I need a girl to pass out the markers" and "Good morning boys and girls"). Although simple statements or requests from a teacher using gender might be interpreted by the child as implying competition, we specifically asked teachers to avoid using gender distinctions for rewards or competitive events. Thus, teachers were asked to limit their use of gender to labeling and classroom organization, and explicitly asked to treat groups equally and to avoid encouraging competition between boys and girls (e.g., we asked teachers to avoid comments such as: "The boys are being quieter so they can line up" or "Let's see who can clean up faster, boys or girls").

Teachers in the low-salience classrooms were given no instructions about changing their behaviors in any way. Importantly, the ongoing policy in both preschools was to avoid using gendered language and organization in the classroom, and thus, children in both conditions had a similar classroom history with the exception of the 2-week period of the study proper. Researchers observed the classrooms every few days over the 2 weeks to observe whether teachers were continuing to use gender (or not) as requested. Indeed, teachers in the high-salience classrooms followed instructions by using gendered language and separating their classroom activities according to gender, whereas teachers in the low-salience classrooms did not organize their classrooms by gender.

Assignment to condition was made at the level of the school rather than at the level of classroom. One reason for this decision was the practical need to meet the administrative needs of the schools, which required having consistent expectations for teachers. Additionally, it carried two design benefits. First, it ensured consistent behaviors among teachers who sometimes move between classrooms or interact with children outside their assigned classroom. Second, it avoided the problem of contamination effects across classrooms, a problem that has been noted in earlier school-based intergroup research (Patterson & Bigler, 2006).

Procedure Overview

Children were interviewed individually both before and after the classroom manipulation with measures described next. In addition and also as explained in more detail later, children were

observed during their normal classroom play periods in sessions at the beginning and end of the classroom manipulation. Following the 2-week period and the administration of all measures, children participated in a debriefing and educational program administered by both teachers and experimenters. This debriefing program was included in an effort to counteract any possible increase in stereotyping and to help children understand prejudice and stereotypes. During the program, puppets, pictures, and posters were used to emphasize similarities and differences among people that were not based on gender. Although we would have liked to have administered delayed assessments of children's gender stereotypes to study long-term effects of the classroom condition and the success of debriefing, practical limitations on the time that the school could devote to the research precluded adding additional assessments to the study.

Gender Attitudes

Children's endorsements of cultural gender stereotypes were assessed before and after the intervention period by the activity and occupation subscales of the Preschool Occupation, Activity, and Trait-Attitude Measure (POAT-AM), developed as a part of a suite of sex-typing measures (see Liben & Bigler, 2002). In this measure, children are shown pictures of activities or occupations, and for each item asked if men/boys, women/girls, or both men and women "should" perform it. Children were asked to respond to 66 occupations and activities, including 22 culturally masculine (e.g., use tools, be a firefighter), 20 culturally feminine (e.g., play with dolls, be a dancer), and 24 neutral (e.g., fly a kite, be a writer) items by pointing to one of three cards to show who they believed should do the activity. The three cards showed schematic (silhouette) pictures of people, with one card showing a man and boy, another showing a woman and girl, and the third containing all of these figures. Scores for this measure are the number of "both men and women" responses. Thus, lower numbers of "both" responses indicate higher endorsement of cultural gender stereotypes and less gender flexibility.

Peer Play

Familiar-peer ratings. To assess children's interest in playing with own- versus other-sex peers at both pre- and posttest, children were shown photo-

graphs of each of their classmates. For each photograph, the child was asked, "How much would you like to play with [name of classmate]?" Children gave their responses by pointing to one of three faces that comprised a 3-point scale. The faces included negative, neutral, and happy expressions, and were explained as meaning they had no interest at all, some interest, or strong interest. Two scores were calculated for each participant, one averaging ratings of children of their own sex, and the other averaging ratings of children of the other sex.

Unfamiliar-peer preference. An additional measure of children's gender-based peer play preferences was administered at posttest only. In this task, children were shown three photographs of groups of unfamiliar peers, one showing all boys, one showing all girls, and the third showing both boys and girls. Each child was told, "These are children from another school," and asked, "If you were to go to that other school, which group of kids would you like to play with?"

Peer play observations. Children were observed during indoor free play periods through a one-way mirror in each classroom. Observations were conducted during free play periods to ensure that teachers were not assigning children to groups in any way. Based on procedures developed by Martin and Fabes (2001), raters observed children sequentially in 15-s units by following a randomly ordered list of children in the classroom. Once raters had observed each child on the list, they started again at the top of the list, until there were three observational periods for each target child on each of 4 days.

Two observers independently noted the number of own- and other-sex peers the child played with during each of the observational periods. There was high reliability between coders on the number of boys and girls with whom the child played (interclass correlations = .96 and .93, respectively). Data from the primary observer were used in the rare cases of disagreement. Two scores were calculated for each period (the beginning and the end of the study): One score was the number of own-sex children and the other score was the number of other-sex children with whom the child played, averaged over the six observations (three observations for each of 2 days at each time period). The pretest observations were calculated from 2 days of observations made just before and as the classroom manipulation was beginning, and the posttest observations were calculated from 2 days of observations made just as and after the classroom manipulation was ending.

Personal Interests

Children's own gender-linked interests were assessed with the activity and occupation subscales of the POAT–Personal Measure (POAT–PM; see Liben & Bigler, 2002). In the POAT–PM, children are asked to indicate the strength of their own interests in various occupations and activities by pointing to one of three faces described earlier for the peer rating measure. Following standard administration procedures, the POAT–PM was administered prior to the POAT–AM to minimize cross-measure contamination.

Instructions were explained with a gender-neutral item, "How much would you like to eat ice cream?" Once children understood the instructions, they were shown the pictures for the POAT–PM proper, and for each item they were asked, "How much would you like to ___?" Responses of *no interest* were scored as 1, *some interest* as 2, and *strong interest* as 3. The dependent measures were calculated separately for masculine and feminine items and were computed as the total number of points awarded divided by the total number of items in that category (i.e., masculine or feminine).

Results

Overview and Preliminary Analyses

Results are described in three major sections that present, first, data on children's gender-stereotyped attitudes; second, data related to children's interest in or observed play with peers of their own or the other sex; and, third, data on children's own interests in stereotypically masculine versus feminine activities and occupations.

Given the relatively narrow range of ages of participants, we had no hypotheses with respect to age differences within our sample. Nevertheless, we conducted preliminary analyses to check for possible age effects both with respect to the way that measures were functioning (e.g., reliability of the POAT scales) and with respect to the effects of the experimental manipulation. The measures were comparably successful for younger and older groups (ages 3.1–4.3 vs. 4.4–5.5, respectively) and effects did not vary with age either in main effects or in interactions. Therefore, results reported are based on the entire sample.

In a second preliminary analysis, we examined correlations between pairs of peer variables to evaluate whether our dependent measures were

tapping redundant constructs. Correlations were uniformly low, suggesting that as intended, the various measures were tapping different aspects of gender-related cognitions and behaviors. Thus, results are presented separately for each measure below.

Gender Attitudes

The dependent variable used to assess gender stereotypes was the total number of "both" (i.e., gender-flexible) responses given to masculine and feminine items on the POAT–AM. Given that the POAT–AM contains 22 masculine and 20 feminine items, scores could range from 0 to 42, with higher scores indicating more flexibility (i.e., lower gender stereotyping). Scores were analyzed with a repeated measures analysis of variance (ANOVA) in which between-subjects variables were condition (low vs. high salience) and participant sex (boys vs. girls), and the within-subjects variable was time (pre- vs. posttest). When appropriate, *t* tests were used as follow-up tests. Differences reported as significant were $p < .05$ or better.

Analyses showed a significant interaction between condition and time, $F(1, 53) = 17.85$, $p < .001$, $\eta_p^2 = .25$. As shown in Figure 1, although in the low-salience condition there was no significant change in gender-stereotyped attitudes from pre- to posttest, $M (SD) = 14.46 (7.14)$ versus

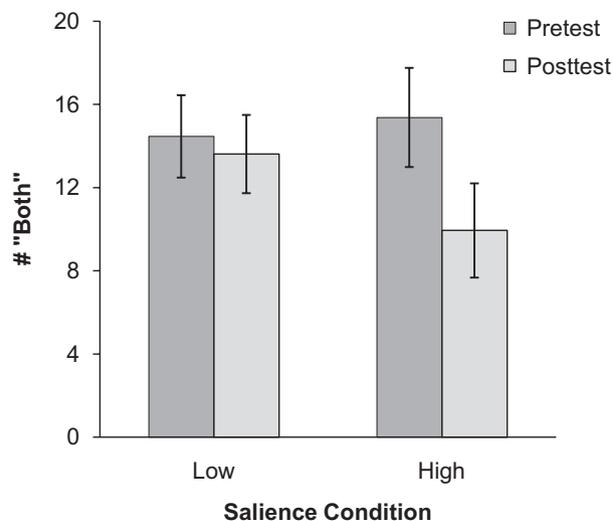


Figure 1. Mean numbers (and standard errors) of "both" (flexible) responses on the POAT–AM gender attitude measure by condition and time.

Note. Higher scores indicate greater gender flexibility or lower gender stereotypes (maximum score = 42).

13.61 (8.13), in the high-salience condition there was a significant drop in the number of flexible "both" responses from pre- to posttest, $M (SD) = 15.37 (6.99)$ versus $9.94 (6.18)$, $d = 0.12$, indicating an increase in gender stereotypes. Also significant was the interaction between time and participant sex, $F(1, 53) = 5.75$, $p = .020$, $\eta_p^2 = .09$. The drop in the number of "both" responses over time was not significant in boys, $14.07 (7.00)$ versus $12.00 (8.00)$, but was significant in girls, $18.89 (6.23)$ versus $13.14 (8.15)$, $d = 0.70$.

Peer Play

Familiar-peer ratings. To examine children's ratings of interest in playing with familiar peers, the average ratings of boy and girl classmates served as the dependent measure in a repeated measures ANOVA. Between-subjects variables were condition (low vs. high salience) and participant sex (boys vs. girls), and within-subjects variables were time (pre- vs. posttest) and peer type (own vs. other sex).

Analyses revealed a significant three-way interaction among condition, time, and peer type, $F(1, 48) = 8.46$, $p = .005$, $\eta_p^2 = .14$, depicted graphically in Figure 2. In the low-salience condition, there was no significant change in other-sex peer ratings between pre- and posttest, $M (SD) = 2.21 (0.47)$ versus $2.36 (0.37)$, whereas in the high-salience condition, there was a significant decrease in other-sex ratings over time, $M (SD) = 2.09 (0.59)$ versus $1.48 (0.36)$, $d = 1.03$. Subsumed by the three-way interaction were two-way interactions between time and condition, $F(1, 48) = 14.26$, $p < .001$, $\eta_p^2 = .22$, and between peer type and condition, $F(1, 48) = 16.39$, $p < .001$, $\eta_p^2 = .26$.

Unfamiliar-peer preference. The dependent variable for assessing interest in playing with unfamiliar peers was the type of group selected, that is, the par-

ticipant's selection of an own-sex group (i.e., the boy-only group for boys or the girl-only group for girls) versus an other-sex group (i.e., either the other- or the mixed-sex group). In both the high- and low-salience groups, only 1% of children chose to play with the other-sex only groups (i.e., those with all boys or all girls). For the purpose of analyses, other- and mixed-sex groups were pooled given that the important question concerned the child's willingness to play with the other sex. Data showed that 37% of children in the low-salience condition chose to play with a group that included children of the other sex compared to only 13% in the high-salience condition, a marginal effect of condition, $\chi^2(1) = 3.71$, $p = .054$.

Peer play observations. To assess children's observed peer play, we calculated average numbers of own- and other-sex peers per observational time unit (15 s). Play with only other-sex and mixed-sex children was collapsed for statistical purposes to cover play with any type of play with other-sex peers. A repeated measures ANOVA was conducted in which between-subjects variables were condition (low- vs. high-salience) and participant sex (boys vs. girls), and within-subjects variables were time (beginning vs. end of 2-week period) and type of peer (own- vs. other-sex).

Analyses revealed a significant three-way interaction among condition, time, and type of peer, $F(1, 47) = 25.46$, $p < .001$, $\eta_p^2 = .35$, shown in Figure 3. As may be seen in the figure, children in the low-salience condition showed no significant change in play with other-sex peers between the beginning and end of the 2-week period, $M (SD) = 0.64 (0.30)$ versus $0.74 (0.39)$, whereas children in the high-salience condition showed a significant decrease in the amount of play with other-sex peers over time $M (SD) = 0.92 (0.41)$ versus $0.26 (0.19)$, $d = 1.3$. Subsumed by the three-way interaction were a marginal two-way interaction between condition and time, $F(1, 47) = 3.90$,

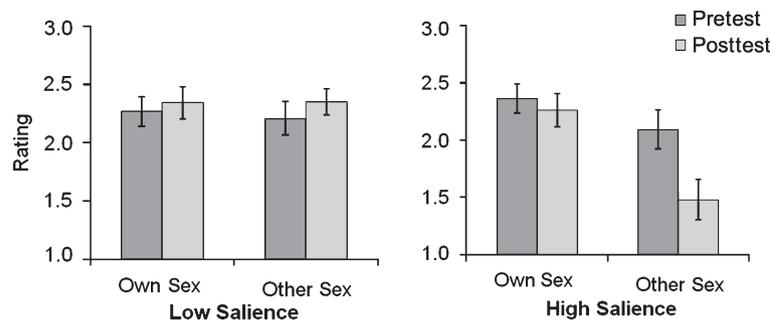


Figure 2. Mean ratings (and standard errors) of children's interest in playing with familiar peers by condition, peer type, and time (most positive rating = 3).

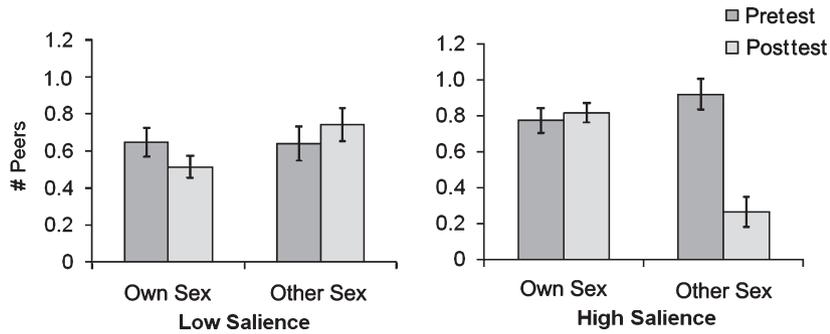


Figure 3. Number of peers (and standard errors) played with, on average, during 15-s observational units by condition, peer type, and time.

$p = .054$, $\eta_p^2 = .08$, and a two-way interaction between time and type of peer, $F(1, 47) = 6.04$, $p = .018$, $\eta_p^2 = .11$.

Personal Interests

As explained earlier, personal interest in culturally gender-stereotyped activities and occupations was assessed by the POAT-PM. Average interest scores (on a 3-point scale) were calculated separately for masculine and feminine items and then analyzed with an ANOVA. Between-subjects variables were condition (low vs. high salience) and participant sex (boys vs. girls), and within-subjects variables were time (pre- vs. posttest) and item type (masculine vs. feminine).

Analyses revealed a significant interaction between participant sex and item type, $F(1, 48) = 205.60$, $p < .001$, $\eta_p^2 = .81$. As expected, boys preferred masculine items more than feminine items, 2.50 (0.24) versus 1.86 (0.28), $d = 2.2$, whereas girls preferred feminine items more than masculine items, 2.48 (0.21) versus 2.11 (0.35), $d = 1.05$. No other main effects or interactions were significant.

Discussion

The major purpose of this research was to test predictions derived from developmental intergroup theory concerning the impact of environmental qualities on the development of stereotypes and prejudices. Specifically, we tested the predictions that making gender more salient in the classroom would lead first, to stronger gender stereotypes and second, to stronger intergroup biases in preschool children. Data provided compelling support for both predictions.

With respect to gender stereotypes, data from the POAT-AM measure showed that children in

classrooms in which gender was made salient expressed significantly more highly stereotyped attitudes at posttest than they had at pretest. A comparable change did not occur among children in the low-salience classrooms, indicating that increased stereotyping in the high-salience condition cannot be attributed merely to retesting effects or to the passage of time.

With respect to in-group versus out-group biases, data from several directions converge on the conclusion that children who experienced high-salience classrooms developed increased rejection or avoidance of the out-group (i.e., the other sex). One line of evidence for this conclusion comes from the familiar-peer ratings. Children in the high- (but not low-) salience condition showed a pre- to posttest decrement in their ratings of how much they would like to play with their other-sex classmates. A second line of evidence comes from the unfamiliar-peer preference measure given at posttest. This measure revealed that children in the high-salience classrooms were marginally ($p = .054$) less likely than children in the low-salience classrooms to say they would like to play with groups of unfamiliar peers that included other-sex children. Finally, a third line of evidence derives from observational data on children's free play in the classroom which revealed that over the 2-week period of the study, children in the high-salience condition showed a significant decrease in their play with other-sex peers whereas children in the low-salience condition showed no such decline.

It is interesting that despite the fact that the measures used could have revealed increased favoritism toward the in-group as well as increased rejection of the out-group, our data showed evidence for the latter only. That is, even in the high-salience group, there were no significant increments over time in children's ratings of own-sex peers or in their actual play with own-sex peers. In earlier

work on race-based social groups, Cameron et al. (2001) found that preschool-aged children held positive in-group bias toward their own racial group, but did not demonstrate negative out-group evaluations, suggesting a primacy for in-group bias. Aboud (2003) found that in-group positivity and out-group negativity were reciprocally correlated for children with little personal experience with out-group members (i.e., from a racially homogeneous school), but in-group positivity and out-group negativity were not reciprocal for children from racially heterogeneous school. In the present study—with gender-defined social groups and coeducational classes—children had experienced extensive daily interactions with other-sex peers, thus offering no reason to expect a reciprocal association. However, it is also risky to assume equivalent processes for groups defined by race and gender, and thus it would be important to continue to explore both in-group and out-group processes and the way that they may be linked in the domain of gender.

Before leaving the discussion of the data relevant to in-group and out-group bias, it is important to point out that when children were tested under the conditions of their normal classroom environment (i.e., at both pre- and posttest for the low-salience condition; at pretest for the high-salience condition), the data showed no evidence of rating own-sex children higher than other-sex children (Figure 2) or of playing more with own-sex children than other-sex children (Figure 3). Given that other investigators have found evidence of own-sex preferences in preschool settings (e.g., Martin & Fabes, 2001), one possible interpretation of the current data is that preschools that assiduously avoid gender-based divisions or language in their classrooms indeed prevent (or postpone) in-group biases observed in other settings. The data may also reflect the egalitarian values of parents who enrolled their children in these schools.

Although more research is thus needed to learn precisely when and how in-group and out-group gender biases evolve, the data from the current study already demonstrate the role of social-group salience on bias, and bolster arguments made by others (Aboud, 2003; Allport, 1954; Brewer, 1999; Cameron et al., 2001) about the importance of assessing in-group and out-group biases with distinct measures.

Taken together, the findings on children's gender stereotypes and group bias extend earlier work with respect to both age and scope of effects of social-group salience. Earlier work (Bigler, 1995)

had demonstrated that increasing gender salience had led to increased stereotypes in older (elementary school) children; the current data demonstrate that these forces operate even among younger children who have less advanced categorization skills. Recent research (Patterson & Bigler, 2006) had shown that highly visible, novel groups that had been defined by colored shirts led to stereotypes and group biases in preschool children. The current data demonstrate that similar effects result from increasing the salience of gender, even though it is already a social category well known to preschoolers (e.g., Ruble, Martin, & Berenbaum, 2006; Weinraub et al., 1984).

As in any study, many important questions remain for future research. One issue concerns how long the observed salience effects persist. As noted earlier, practical constraints prevented the inclusion of delayed posttests in this study. However, having now established immediate and dramatic effects of gender salience in the classroom, it becomes important for future research to address the degree to which such effects are long lasting, and to identify the factors that influence their longevity.

Another question concerns the associations among the outcome variables. We had selected a range of measures addressed to constructs hypothesized to be affected by the gender-salience manipulation. One measure (POAT-AM) was selected to assess the degree to which the child endorsed cultural gender-stereotyped attitudes about what others should do. Two measures (peer ratings and, at posttest only, the novel peer measure) were selected to assess children's cognitions about their interest in playing with others of their own and the other sex. One measure (play observation) was selected to study children's actual behavior with their own- and other-sex peers. All measures for which there were pre- and posttest data showed a significant impact of the manipulated classroom condition, and the posttest only novel peer measure showed a marginal effect of group. Thus, the data from the current study allow us to conclude that as hypothesized, increasing gender salience in the classroom affects a range of children's gender-related beliefs and behaviors. Remaining for later research with much larger samples is the intriguing and important question about how the various components of gender beliefs and behaviors are connected, in some cases, perhaps causally. For example, future research might examine whether a rise in gender stereotyping decreases children's willingness to play with children of the other sex, or whether there is an inverse pathway, or if

perhaps each outcome is instead controlled by some other factor.

A clear implication of the findings already in hand is that schools should make it as unacceptable to use gender-specific language and divisions (e.g., "Good morning boys and girls") as it is to use race-specific language and divisions (e.g., "Good morning Black children and White children"). There are ample anecdotal data that the former continues to occur in our educational system. For example, as this article was being written, the second author received a note from a former student reporting that she had worked over the summer in a preschool center "and they actually split the children up for free play time, saying girls you can play this and boys you can play that, i.e., dolls and action figures" (D. Russell, personal communication, October 7, 2008).

What also remains for future research is exploration of the cognitive mechanisms by which increased gender salience leads to greater gender stereotypes and out-group biases. As proposed by Bigler and Liben (2006, 2007), developmental intergroup theory posits that one causal mechanism for stereotyping lies in children's active attempts to assign meaning and correlates to social categories that are made salient in the environment. To test this position, it would be necessary to monitor children's cognitive processing (e.g., measure children's attention to implicit and explicit correlates of gender) over time under low- versus high-salience conditions.

Another potential causal mechanism involves the peer environment. Previous work has shown that peer groups provide a powerful context to socialize gender-linked play styles and preferences (Fabes, Martin, & Hanish, 2003; Maccoby & Jacklin, 1987). In a longitudinal study in preschool classrooms, Martin and Fabes (2001) found that children who were engaged in high levels of own-sex play at the beginning of the study showed more gender-typed play at the end of the study than did children who were engaged in low levels of own-sex play. They noted that over time, the play styles of each gender group became increasingly distinct.

Several aspects of our work showed that the high-gender-salience condition led children to reject other-sex peers, thereby increasing gender segregation. First, anecdotes reported by teachers in the high-salience condition suggest that children were quick to pick up on gender categorization and organize themselves by gender. One teacher, for example, noted that on the second day of the manipulation, without prompting from

her, children established separate boys' and girls' snack tables. Second, the formal observational data on peer play showed that children in the high-salience condition became less likely to play with children of the other sex over time. To the degree that increased gender salience in the classroom led children to play less with members of the out-group, children would have even less occasion to be socialized into the play styles and preferences typical of the other sex, in turn exaggerating differences between gender groups. It is likely, given previous findings and our current study, that early gender segregation leads to two distinct gendered peer cultures, which, exacerbated over time, establish a script for later relationship styles and affect life decisions (Leaper, 1994).

In short, anecdotal data on children's classroom behaviors and systematic data on children's peer play suggest that much of the effect of the teachers' greater use of gender in the high-salience groups may have had its impact by inspiring children themselves to create a highly gendered environment. Another important direction for future research is therefore to observe, in detail, how the classroom environment (e.g., children's communication, segregation, and role playing) evolves under low- versus high-salience conditions.

Finally, the data from the current research add to the literature concerning the strong and persistent nature of children's own gendered interests. As expected, responses to the POAT-PM showed that children held sex-linked preferences, with boys preferring masculine items more than feminine items and girls preferring feminine items more than masculine items. There was no modification of this pattern as a function of gender-salience condition. The finding is consistent with earlier research (e.g., Bigler, 1995; Bigler & Liben, 1990) showing that brief experimental interventions leave children's own gendered interests and preferences unchanged even in the face of significant changes in children's beliefs about what others should do. Such findings suggest that modifying children's gender stereotypes will not be sufficient to expand children's educational and occupational choices (e.g., encouraging more girls and women to pursue science-related hobbies, coursework, and careers) and that more directed intervention strategies are needed. It is plausible that the 2-week intervention was not long enough to affect children's own interests and that over a longer period of time, children would be likely to internalize the gender stereotypes and show more gender-linked interests.

In conclusion, the findings from the present research have important implications for both theory and practice. At the theoretical level, data provide support for a core prediction of developmental intergroup theory (Bigler & Liben, 2006, 2007), which posits that social-group salience plays an important role in the formation of stereotypes and biases. At the practical level, data demonstrate that there is a pervasive, powerful, and remarkably quick effect of making gender salient in the classroom. The current study highlights the impact in one area of a child's environment (the classroom), but one would also expect a strong influence of the attention to social-group categories in the home (e.g., different chores for male and female members of the household) and in the broader culture (e.g., different professional sports leagues) as well. Although there is much to learn about the mechanisms by which increasing the salience of group membership operates, the current research demonstrates the importance of minimizing attention to social-group divisions in young children's classroom environments, even for social-group divisions, like gender, that are already visible and salient in the society at large.

References

- About, F. E. (1988). *Children and prejudice*. New York: Blackwell.
- About, F. E. (2003). The formation of in-group favoritism and out-group prejudice in young children: Are they distinct attitudes? *Developmental Psychology, 39*, 48–60.
- About, F. E., & Amato, M. (2001). Developmental and socialization influences on intergroup bias. In G. J. O. Fletcher & M. S. Clark (Series Eds.) & R. Brown & S. Gaertner (Vol. Eds.), *Blackwell handbook in social psychology: Vol. 4. Intergroup processes* (pp. 65–85). Oxford, UK: Blackwell.
- Abrams, D., Rutland, A., & Cameron, L. (2003). The development of subjective group dynamics: Children's judgments of normative and deviant in-group and out-group individuals. *Child Development, 74*, 1840–1856.
- Allport, G. (1954). *The nature of prejudice*. Reading, MA: Addison-Wesley.
- Bem, S. L. (1983). Gender schema theory and its implications for child development: Raising gender-aschematic children in a gender-schematic society. *Signs: Journal of Women in Culture and Society, 8*, 598–616.
- Bigler, R. S. (1995). The role of classification skill in moderating environmental influences on children's gender stereotyping: A study of the functional use of gender in the classroom. *Child Development, 66*, 1072–1087.
- Bigler, R. S., Jones, L. C., & Lobliner, D. B. (1997). Social categorization and the formation of intergroup attitudes in children. *Child Development, 68*, 530–543.
- Bigler, R. S., & Liben, L. S. (1990). The role of attitudes and intervention in gender-schematic processing. *Child Development, 61*, 1440–1452.
- Bigler, R. S., & Liben, L. S. (2006). A developmental intergroup theory of social stereotypes and prejudice. In R. V. Kail (Ed.), *Advances in child development and behavior* (Vol. 34, pp. 39–89). San Diego, CA: Elsevier.
- Bigler, R. S., & Liben, L. S. (2007). Developmental intergroup theory: Explaining and reducing children's social stereotypes and prejudice. *Current Directions in Psychological Science, 16*, 162–171.
- Brewer, M. B. (1999). The psychology of prejudice: Ingroup love or outgroup hate? *Journal of Social Issues, 55*, 429–444.
- Cameron, J. A., Alvarez, J. M., Ruble, D. N., & Fuligni, A. J. (2001). Children's lay theories about ingroups and outgroups: Reconceptualizing research on prejudice. *Personality and Social Psychology Review, 5*, 118–128.
- Fabes, R. A., Martin, C. L., & Hanish, L. D. (2003). Young children's play qualities in same-, other-, and mixed-sex peer groups. *Child Development, 74*, 921–932.
- Gelman, S. A. (2003). *The essential child*. New York: Oxford University Press.
- Grusec, J. E., & Brinker, D. B. (1972). Reinforcement for imitation as a social learning determinant with implications for sex-role development. *Journal of Personality and Social Psychology, 21*, 149–158.
- Killen, M., Margie, N. G., & Sinno, S. (2006). Morality in the context of intergroup relations. In M. Killen & J. Smetana (Eds.), *Handbook of moral development* (pp. 155–184). Mahwah, NJ: Erlbaum.
- Killen, M., & McKown, C. (2005). How integrative approaches to intergroup attitudes advance the field. *Journal of Applied Developmental Psychology, 26*, 616–622.
- Koblinsky, S. A., Cruse, D. F., & Sugawara, A. I. (1978). Sex-role stereotypes and children's memory for story content. *Child Development, 49*, 452–458.
- Kowalski, K. (2003). The emergence of ethnic and racial attitudes in preschool-aged children. *Journal of Social Psychology, 143*, 677–690.
- Leeper, C. (1994). Exploring the consequences of gender segregation on social relationships. *New Directions for Child Development, 65*, 67–86.
- Liben, L. S., & Bigler, R. S. (2002). The developmental course of gender differentiation: Conceptualizing, measuring, and evaluating constructs and pathways. *Monographs of the Society for Research in Child Development, 67*(2, Serial No. 269).
- Liben, L. S., & Signorella, M. L. (1980). Gender-related schemata and constructive memory in children. *Child Development, 51*, 11–18.
- Maccoby, E. E., & Jacklin, C. N. (1987). Gender segregation in childhood. In H. W. Reese (Ed.), *Advances in child development and behavior* (Vol. 20, pp. 195–216). New York: Academic Press.

- Martin, C. L., Eisenbud, L., & Rose, H. (1995). Children's gender-based reasoning about toys. *Child Development, 66*, 1453–1471.
- Martin, C. L., & Fabes, R. A. (2001). The stability and consequences of young children's same-sex peer interactions. *Developmental Psychology, 37*, 431–446.
- Martin, C. L., & Halverson, C. F. (1981). A schematic-processing model of sex typing and stereotyping in children. *Child Development, 52*, 1119–1134.
- Nesdale, D. (2004). Social identity processes and children's ethnic prejudice. In M. Bennett & F. Sani (Eds.), *The development of the social self* (pp. 219–245). New York: Psychology Press.
- Patterson, M. M., & Bigler, R. S. (2006). Preschool children's attention to environmental messages about groups: Social categorization and the origins of intergroup bias. *Child Development, 77*, 847–860.
- Ruble, D. N., Martin, C. L., & Berenbaum, S. A. (2006). Gender development. In W. Damon & R. M. Lerner (Series Ed.) & N. Eisenberg (Vol. Ed.) *Handbook of child psychology* (6th ed., Vol. 3, pp. 858–932). New York: Wiley.
- Rutland, A. (1999). The development of national prejudice, in-group favoritism and self-stereotypes in British children. *British Journal of Social Psychology, 38*, 55–70.
- Rutland, A., Cameron, L., Milne, A., & McGeorge, P. (2005). Social norms and self-presentation: Children's implicit and explicit intergroup attitudes. *Child Development, 76*, 451–466.
- Sherif, M., Harvey, O. J., White, B. J., Hood, W. R., & Sherif, C. W. (1961). *Intergroup conflict and cooperation: The Robbers Cave experiment*. Norman, OK: University Book Exchange.
- Signorella, M. L., Bigler, R. S., & Liben, L. S. (1997). A meta-analysis of children's memories for own-sex and other-sex information. *Journal of Applied Developmental Psychology, 18*, 429–445.
- Stern, M., & Karraker, K. H. (1989). Sex stereotyping of infants: A review of gender labeling studies. *Sex Roles, 20*, 501–522.
- Tajfel, H., & Turner, J. C. (1986). The social identity theory of intergroup behavior. In S. Worchel & W. G. Austin (Eds.), *Psychology of intergroup relations* (2nd ed., pp. 7–24). Chicago: Nelson-Hall.
- Waxman, S., & Gelman, R. (1986). Preschoolers' use of superordinate relations in classification and language. *Cognitive Development, 1*, 139–156.
- Waxman, S. R., & Markow, D. B. (1998). Object properties and object kind: Twenty-one-month old infants' extension of novel adjectives. *Child Development, 69*, 1313–1329.
- Weinraub, M., Clemens, L. P., Sockloff, A., Ethridge, R., Gracely, E., & Myers, B. (1984). The development of sex role stereotypes in the third year: Relationships to gender labeling, gender identity, sex-typed toy preferences, and family characteristics. *Child Development, 55*, 1493–1503.
- Yee, M. D., & Brown, R. (1992). Self-evaluations and intergroup attitudes in children aged three to nine. *Child Development, 63*, 619–629.
- Yee, M. D., & Brown, R. (1994). The development of gender differentiation in young children. *British Journal of Social Psychology, 33*, 183–196.