

## Children's Social Reasoning about Inclusion and Exclusion in Gender and Race Peer Group Contexts

*Melanie Killen and Charles Stangor*

This study investigated whether children's and adolescents' judgments about exclusion of peers from peer group activities on the basis of their gender and race would differ by both age level and the context in which the exclusion occurred. Individual interviews about exclusion in several different contexts were conducted with 130 middle-class, European American children and adolescents. Younger children were expected to reject exclusion, by using judgments based on moral reasoning, regardless of the potential cost to group functioning, whereas older children were expected to condone exclusion on the basis of group membership in cases in which the inclusion of these children might interrupt effective group functioning. On measures of judgments, justifications for those judgments, and ratings of the appropriateness of exclusion, the vast majority of children used moral reasoning and rejected exclusion in contexts in which only the presence of a stereotype justified it. As expected, however, older children (13 years) were more likely to allow exclusion than younger children (7 and 10 years) when group functioning was threatened, and they justified this exclusion by using appeals to effective group functioning.

### INTRODUCTION

Many studies have shown that children, as early as the preschool years, are well aware that it is wrong to harm and to act in ways that are unfair to others (for reviews, see Killen, 1991; Smetana, 1995; Tisak, 1995; Turiel, Killen, & Helwig, 1987). These studies, however, have primarily focused upon direct physical harm (such as hitting) or denial of access to resources (such as unfair distribution; see Turiel, 1983, 1998). Much less is known about whether children view other types of potentially harmful activities, such as excluding children from peer group activities on the basis of group memberships, as unfair (but see Short, 1993; Theimer, Killen, & Stangor, 2001). Although the literature on peer exclusion (in general) is quite extensive, most of this work has concentrated on the traits of the individual child being excluded and how this accounts for peer rejection (Asher & Coie, 1990). Recently, one study has examined the behavioral patterns of exclusion by groups (see Zarbatany, Van Brunschot, Meadows, & Pepper, 1996), but no work, to the author's knowledge, has examined how children evaluate peer exclusion from the viewpoint of the group and intergroup relationships (an exception is Theimer et al., 2001). Understanding how children reason about such decisions, however, is important because decisions about the appropriateness of inclusion or exclusion represent an integral part of reasoning about intergroup relationships in adults (such as including or excluding women from the military services or minorities from job opportunities; e.g., Macrae, Stangor, & Hewstone, 1996).

One study that has investigated this issue has shown that young children judge it as wrong to exclude someone from a play activity solely on the basis of their gender. Theimer et al. (2001) found that preschool children thought it was wrong for a group of girls to exclude a boy from playing with dolls, even though these same children saw doll-playing as an activity that was primarily performed by girls (likewise, it was viewed as wrong for a group of boys to exclude a girl from playing with trucks, even though truck-playing was seen as an activity more appropriate for boys). The goal of the present research was to study the developmental trajectory of decisions about inclusion and exclusion in gender and race peer contexts into adolescence.

On the basis of social-cognitive development theory, it was hypothesized that decisions about the appropriateness of excluding children from social groups (particularly, gender and race) involve two forms of social reasoning—moral beliefs about the wrongfulness of exclusion, and social-conventional beliefs about social group processes and group functioning. Moral beliefs include concepts about fairness and rights, equal treatment, and equal access (Damon, 1983; Turiel, 1998). Social-conventional beliefs entail several forms of reasoning, including those that concern group functioning (Turiel, 1978, 1983, 1998), group identity (Brown, 1989), and stereotypes about others based on their group membership (Carter &

Patterson, 1982; Liben & Signorella, 1993; Stangor & Ruble, 1989; Stoddart & Turiel, 1985).

Research on children's reasoning about social conventions has shown that social-conventional concepts change with age, and particularly so in terms of taking social group roles and expectations into account (Helwig, 1995, 1997; Killen, 1991; Turiel, 1978, 1983, 1998). Whereas young children reason about social conventions in terms of social uniformity and rule systems (e.g., "It's wrong to call a teacher by her first name because there is a rule about it"), older children reason about social group customs in terms of societal standards and social coordination (e.g., "It's wrong to call a teacher by her first name because maybe the other students would think of her as a peer instead of someone with authority and a higher status"; see Turiel, 1983, p. 103). With age, children become increasingly concerned about the nature of social groups, the norms and expectations that go along with the structure of the group, and effective group functioning.

Theoretically, then, decisions about potential exclusion from social groups involve the coordination of moral judgments about the wrongfulness of exclusion with a range of social-conventional judgments about social group functioning, group identity, and group stereotypes. Evaluating acts of exclusion from groups involves weighing these competing moral and social-conventional considerations. Given that research has shown that children are sensitive to context issues (Helwig, 1995) and that, with age, children increasingly use more complex social group reasoning processes (Tisak, 1995; Turiel, 1983), it was predicted that children's judgments about exclusion from peer groups would be sensitive to the context in which the exclusion occurred and that sensitivity to the impact of exclusion on effective group functioning would increase with age.

To test these hypotheses, children were asked to make judgments about the appropriateness of exclusion in gender and racial peer group contexts. Gender and race were selected for several reasons. First, gender and race are two of the most salient social group membership categories to emerge in development (see Aboud, 1992). Second, gender and racial stereotypes are a common source of prejudice and contribute to intergroup tensions and conflicts. In addition, children's and adolescents' exclusion in peer group contexts is most frequently about gender (Maccoby, 1988) and race (Aboud, 1992). No studies of which the authors are aware have compared children's judgments about gender and racial exclusion in peer group contexts, but research on stereotypes indicates that both gender and racial stereotypes emerge during the preschool period (Aboud, 1992).

Three exclusion contexts were designed for children's evaluation. In the first context, called the *straightforward exclusion context*, a group of peers is considering excluding a child from the group for solely stereotypic reasons, without any other justification except that the children might feel "uncomfortable" having the child who was not stereotypical for the activity (e.g., a boy in a ballet class) in the group. It was expected that despite the possibility of using the stereotype to justify exclusion, moral reasons would prevail for all children in this context. Straightforward exclusion contexts are similar to straightforward moral transgressions in which one person inflicts harm on another for no reason and there are few competing considerations.

Two *multifaceted* contexts were also created. In these, the cost to group functioning was increased and the morally relevant salience of the exclusion decision was decreased. In each of these contexts, the participant was asked to pick one of two children to join the group—one who fit the stereotype of the peer group activity and one who did not fit the stereotype. This resulted in the inclusion of one child and the exclusion of another, which made the decision to exclude a child who did not fit the stereotype less morally difficult because one was including someone else at the same time.

The cost to group functioning was manipulated by providing information about the child's qualifications to join the group between the two multifaceted contexts. In the first context, called *equal qualifications*, two children were said to be equally qualified to join the group (e.g., "A boy and a girl are equally good at ballet"). In the second context, called *unequal qualifications*, the child who did not fit the stereotype was said to be less qualified than the child who fit the stereotype (e.g., "The girl is better at ballet than is the boy"). The child who did not fit the stereotype was made less qualified to test the extent to which children would continue to apply their judgment about not excluding others on the basis of gender and race (as measured in the straightforward context) in different contexts with competing considerations. Would children who judged it wrong to exclude solely on the basis of gender or race (straightforward context) be willing to exclude a child when that child was less qualified than another child who fit the stereotype of the activity?

It was expected that both moral and social-conventional reasoning would be relevant to decisions in multifaceted contexts. In these contexts, it was expected that decisions to include the child who did not fit the stereotype would be justified in terms of moral reasons of fairness and equity, whereas decisions to

exclude the child who did fit the stereotype would be justified in terms of social conventional reasons regarding effective group functioning. On the one hand, moral reasoning would be reflected when it was judged as wrong to exclude someone who did not have an equal opportunity to join (e.g., fair and equal treatment). On the other hand, social conventional reasoning would be reflected in children's reasoning when it was judged as all right to exclude a child on the basis of social-conventions, such as stereotypes (e.g., "Ballet is for girls"; see Carter & Patterson, 1982; Stoddart & Turiel, 1985), group functioning (e.g., "The group will work better with someone who knows how to do the activity"), or group identity (e.g., "The club needs to feel like a group"). It was expected that, with age, children would increasingly focus on group functioning and view it as wrong to exclude someone who did not have familiarity with the activity and thus could not contribute to the group identity or the effective functioning of the group.

Further, it was expected that, with age, children would be increasingly able to differentiate between different contexts of exclusion. The difference between the equal and the unequal qualifications contexts is that in the unequal qualifications context (in which the girl is better at ballet, for example), the cost of including the child for whom the activity is not stereotypical (such as a boy for ballet) is higher than in the equal qualifications context. In the equal qualifications context both children are equally qualified at the activity (even though one child is more typically associated with the activity). It was expected that children would be more likely to choose to include the nonstereotypical child in the equal qualifications context than in the unequal qualifications context. The reasons for picking the nonstereotypical child in the equal qualifications context could be due to either judgments about equal access or to judgments about group functioning. It was predicted that it would be more likely for equal access justifications to emerge in the equal qualifications context than in the unequal qualifications context given the higher cost to group functioning in the unequal qualifications context (when one child is less qualified to participate in the peer group activity). It was also expected that older children would be more likely than younger children to differentiate their judgments by context because older children would be more sensitive to the potential influence upon effective group functioning in the unequal qualifications context.

To test these hypotheses, European American children, equally divided between boys and girls, between the ages of 7 and 13 were interviewed regarding their beliefs about the appropriateness of excluding

children from stereotypical peer group activities on the basis of their gender and race.<sup>1</sup>

Four different peer group activities, stereotypical of girls, of boys, of White children, and of Black children, were used. These activities were ballet club, baseball card club, math club, and basketball club, respectively. Although three peer group activities associated with gender or race categories (e.g., ballet for girls, baseball card clubs for boys, basketball for Black children) were identified, finding a peer group activity associated with White children was more difficult. A math club was found to be somewhat associated with White children, particularly for older children. Thus, the math club was used even though it was somewhat different from the other three clubs because of the academic content. It was emphasized that the clubs were voluntary and not associated with school work. After-school peer group activities were chosen so that children's evaluations of authority would not enter into their decision making. To avoid introducing racial stereotypes to younger children, the words "Black" and "White" were not mentioned to children in the race stories. Instead, the picture cards were simply shown and the child was asked whether it would be all right or not all right to exclude the child "standing at the door?" The children rated the appropriateness of excluding children from these peer group activities for each of the three separate exclusion contexts. In addition, a group activity knowledge assessment was administered to ensure that the participants viewed the peer group activities as associated with gender and race group membership.

The research design also allowed us to test two subsidiary hypotheses. In the Theimer et al. (2001) study on preschool-aged children's evaluations, it was found that girls were more likely to evaluate exclusion negatively than were boys. In general, several studies have indicated that girls are more sensitive to prosocial issues and to exclusion than are boys (Killen & Turiel, 1998; Wentzel & Erdley, 1993; Zahn-Waxler, Cole, Welsh, & Fox, 1995; Zarbatany et al., 1996). On the basis of these findings it was predicted that girls would rate exclusion more harshly across all contexts.

<sup>1</sup> All children whose parents gave consent were interviewed. In addition to the 130 European American students, 31 African American students were also interviewed. Originally, the study was designed to include an equal number of European American and African American students to include race as a participant variable. Because an equal race sample size was not obtained, only the findings for the European American sample are reported in this paper. Results from the sample of African American children do, however, provide some information about how the findings might generalize to other groups, and this is mentioned briefly in the Discussion section. A larger African American sample is currently being collected.

Second, research with children and adults has indicated that individuals often display an ingroup bias or *ingroup favoritism* (Bennett, Barrett, Lyons, & Sani, 1998; Brewer, 1979; Damon, 1977, chapter 3; Mackie, Hamilton, Susskind, & Rosselli, 1996; Tajfel, Billig, Bundy, & Flament, 1971; Van Avermaet & McClintock, 1988; Yee & Brown, 1992), in which they make more positive judgments or assign more positive rewards to other members of their own group. No research, however, has been conducted to determine whether children display an ingroup bias in their social judgments about exclusion. In this study, whether girls and boys were more willing to condone exclusion of opposite-sex than same-sex children from group activities was tested. Finally, to provide breadth for the types of social group categories children think about, exclusion scenarios were included for both gender and race. No hypotheses were formulated to distinguish these types of exclusion, however, because of a lack of prior findings directly bearing on this comparison.

## METHOD

### Participants

Participants were 65 girls and 65 boys from three grades. There were 19 female and 20 male first graders,  $M = 6.6$ ,  $SD = .4$ , *range* = 5.11–7.4, 25 female and 23 male fourth graders,  $M = 9.6$ ,  $SD = .5$ , *range* = 8.11–10.9, and 21 female and 22 male seventh graders,  $M = 12.6$ ,  $SD = .4$ , *range* = 12.0–13.5. All students were European American. The children were enrolled in mixed-ethnicity, middle-class public schools in a suburban area of Maryland, outside of Washington, DC. All students were informed that the interviews were confidential, voluntary, and anonymous. Parental consent was obtained for all participants.

### Procedure and Design

All students were individually interviewed for about 35 min by a graduate research assistant in a quiet room at school. The three parts of the interview, a warm-up task, a group activity knowledge assessment, and a group exclusion evaluation, were administered to all children.

*Warm-up task.* The purpose of the warm-up task was to familiarize students with the use of the 7-point Likert response format that was to be used in the group exclusion evaluations and to validate that the students could use the scale. Three transgressions were described, accompanied by picture cards, and participants were asked to rate the "badness" of each

of three acts by using a scale (0 = not at all bad; 6 = very, very bad). The three acts were hitting someone for no reason, calling a principal by her first name, and sleeping late. Confirming that the children understood the task and were able to use the Likert scale measure, the three acts were rated significantly differently,  $Ms = 5.5$ , 2.5, and 2.14, respectively, in a manner that accurately reflects the severity of the transgressions (there were no significant age or gender differences).

*Group activity knowledge assessment.* In the knowledge assessment task, children were asked to decide who liked to do particular activities by pointing to a laminated  $8\frac{1}{2}'' \times 11''$  card that had a row of five sets of very simple "smiley" faces. To depict girl as opposed to boy faces, a bow was drawn on the head of the girls; to depict White versus Black faces, the Black faces were shaded. To assess knowledge of gender-related activities, children were asked "Who likes to do X?" and were asked to point to one of five sets of faces: (1) only girls: four girl faces; (2) mostly girls: one boy and three girl faces; (3) same: two boy and two girl faces; (4) mostly boys: three boy faces and one girl face; (5) only boys: four boy faces. The questions were ballet (girl activity), baseball cards (boy activity), tennis (neutral activity), and reading books (neutral activity). Responses were coded on a 5-point scale corresponding to each choice (1 = only girls to 5 = only boys). For the race measure, children were asked "Who likes to do X?" and were asked to point to one of five sets of faces: (1) only Black children: four Black child faces; (2) mostly Black children: one White child face and three Black child faces; (3) same: two Black and two White child faces; (4) mostly White children: three White child faces and one Black child face; and (5) only White children: four White child faces. The questions were math (White activity), basketball (Black activity), art (neutral activity), and singing (neutral activity). Responses were coded on a 5-point scale corresponding to each choice (1 = only Black children to 5 = only White children). The order in which the gender and race items were presented was counterbalanced.

*Group exclusion evaluation.* The group exclusion evaluation comprised four descriptions of children in different after school peer clubs in which the group was considering excluding a child from participating in the group. The exclusions involved ballet (girls exclude a boy), baseball cards (boys exclude a girl), math (White children exclude a Black child), and basketball (Black children exclude a White child). Each exclusion event was described by using  $8\frac{1}{2}'' \times 11''$  picture cards that illustrated the scene (e.g., for the ballet scenario, there was a picture of girls in a room

with a ballet bar and a mirror and a boy at the door looking in the room). For the race contexts, the gender of the children described in the event matched the gender of the participant. The interviewer began the interview by saying, "I am going to tell you about a number of different afterschool clubs and some of the things that happened to the kids in the clubs. These clubs are for kids and there are no teachers or adults in the clubs. These kids are about your age. There are no right or wrong answers. I'm just interested in whatever you think about these stories."

For each of the four exclusion events, three different judgmental contexts were presented to the child—straightforward exclusion, equal qualifications, and unequal qualifications. Furthermore, in describing each of the contexts, the children were told that some children were of one opinion about the appropriateness of exclusion, whereas other children had the opposite opinion. This was done to suggest that either of the two decisions would be appropriate because some children already favored it. These competing suggestions were made to avoid problems related to social desirability (see Harter, 1998).

In the straightforward exclusion contexts, participants were told that a child wanted to join an activity and it was said that some of the children in the group would be uncomfortable if the child joined and that they might quit if the child was included—that is, that some children would like the (peer) to be included and some would not like the (peer) to be included. In the equal qualifications context, participants were told that two children wanted to join the club (one child fit the stereotype and one child did not), that the two children were known to be equally good at the activity, and that the club members were divided about who to include—some of the children want to give the (nonstereotypical peer) a chance because (he or she) does not usually do the activity, whereas the others think that it would be good to have another (peer) join who was like the others. Finally, in the unequal qualifications context, the children were again told that two children wanted to join the club—one child fit the stereotype and one child did not, that one child was better at the activity (than the one who fit the stereotype), and (again) that the children were divided about which child to include.

As an example, in the ballet scenario, for the straightforward exclusion context, participants were told that some of the girls in the ballet club did not want a boy to join because they would be uncomfortable. Children were asked whether it was alright or not alright to let the boy into the ballet club. In the equal qualifications context, participants were told that both a boy and a girl wanted to join the club but

there was room for only one more person to join and the boy and girl were equally good at ballet. In the unequal qualifications context, participants were told that two children, a boy and a girl, both wanted to join the ballet club but that there was only room for one more person and that the girl was better at ballet.

Five assessments were made of participants' responses. First, participants were asked for their *judgment* about the exclusion in the straightforward context: "Is it alright or not alright for the group to exclude the [nonstereotypical] child from the activity?" Then they were asked for their *choice* of who to pick (in both the equal qualifications and in the unequal qualifications scenarios, e.g., "Who should the club pick?"). Children were asked for their *justifications* about their judgments (straightforward context) and choices (multifaceted contexts). Children were also asked how bad it would be for the club members to exclude the child who was not stereotypical for the activity and to make a *rating* of how bad excluding this child would be on a scale from 0 (Not at all bad) to 6 (Very, very bad). Because it had been found in pilot testing that questions about race were more socially sensitive than questions about gender, the gender scenarios were always described first, followed by the race scenarios; and girls heard the girl-excluded scenario first whereas boys heard the boy-excluded scenario first. In addition, a separate analysis was conducted on the number of times that children invoked explicit *stereotypes* about the group activity when providing a justification for their judgments and choices (for instance, "Boys aren't good at ballet"). Because this analysis revealed that very few of the children (less than 8%) used explicit stereotypes to justify exclusion, no further analyses were conducted for this assessment.

### Coding and Reliability

Judgments and choices were coded dichotomously. Justifications were coded by using a coding system based both on previous categories used in the literature (Smetana, 1995; Tisak, 1995) and on the results of pilot data. The coding category system comprised three moral (fairness, equal treatment, equal access) and three social-conventional (social conventions, group functioning, and group identity) codes (see Table 1). For the primary analyses, the three moral categories were collapsed as "moral" and the three social-conventional were collapsed as "social-conventional." Three subcategories, included in the pilot system, were deleted in the final version of the coding system because of low frequency (less than 6%). Two moral subcategories were deleted: prosocial ("You

**Table 1** Justification Coding Categories**Moral**

*Fairness and rights.* Appeals to the maintenance of fairness in the treatment of persons (e.g., "It wouldn't be fair to exclude him"), to the rights of individuals (e.g., "She has a right to join the club if she wants to"), and to the wrongfulness of discrimination based on race or gender or both (e.g., "You shouldn't discriminate against someone just because of their color or gender").

*Equal treatment.* Appeals to the equal treatment of individuals (e.g., "Everyone should be treated the same").

*Equal access.* Appeals to the learning opportunities of those, as members of discriminated groups, who have not previously had the chance (e.g., "Boys should have a chance to do ballet because they usually don't get to do it"; "Teach her about baseball cards because girls don't often get a chance and they should have the same opportunity").

**Social Conventional**

*Social conventions.* Appeals to the expectations of the group (e.g., "The other kids would think that John is strange if he takes ballet"; "The boy would feel strange if he was in a group with all girls"), traditions, customs, and norms of a society (e.g., "Girls are not supposed to play with baseball cards"; "Girls usually like to play jump rope and games like that").

*Group functioning.* Appeals to making the group function well. This includes statements about admitting someone who will make the club more enjoyable or more interesting to its members (e.g., "Admit the one who is more qualified because then the club will know more and work much better as a group together"; "Choose her because she's better at ballet and that's what their club's about, and so you want to have the better person, because they'll probably enjoy it more"; "Since he knows a lot more he can be more useful in the club"; "Because they don't want someone in there who doesn't know the same as them. They might get bored of . . . like when they say something she might not understand and they have to explain it.")

*Group identity.* Appeals to the identity of the group (e.g., "The black kids on the basketball team need to have their own team"; "The girls will feel uncomfortable if a boy is in the club") and group decision making ("The group can decide whatever they want").

should include someone in order to be nice"), from research by Eisenberg and Miller (1987); and individual merit ("A person who is good at something deserves to be in the club"), from research by Damon (1977); and a third social-conventional subcategory was deleted: stereotypic beliefs ("Boys are not good at ballet"), from research by Carter and Patterson (1982).

Responses that used more than one type of justification were coded for each applicable justification, although less than 5% of the participants used more than one justification per response. Reliability was conducted on 25% of the protocols (390 data points) by two trained coders. For justifications in the straightforward context, Cohen's  $\kappa = .88$ , for the

equal qualifications context,  $\kappa = .80$ , and for the unequal qualifications context,  $\kappa = .88$ .

## RESULTS

The report of the analyses begins with the findings for the Group Activity Knowledge measure to demonstrate that children were aware of the normatively appropriate activities for the target groups that we were using. Next, the analyses for judgments, ratings and justifications within the straightforward exclusion context is described to affirm that all children at all ages saw such unjustified exclusion as wrong. Then the two multifaceted contexts (equal qualifications and unequal qualifications) are compared. Analyses were conducted on three separate measures—judgments, ratings, and justifications. Judgments were coded dichotomously (0 and 1), ratings were on a 7-point scale, and justifications were the proportion of moral and social conventional justifications. The findings for our expectations that inclusion would be justified on moral grounds, whereas exclusion would be based on social conventional reasoning, are also described. In all subsequent analyses, post hoc comparisons were performed by using Bonferroni corrections to correct for Type I error. In cases where proportions were used, arcsine transformations were conducted to normalize the distributions (Winer, 1971).

### Group Activity Knowledge Assessment

To investigate whether children held the appropriate beliefs about the gender activities, 2 (gender of child)  $\times$  3 (grade)  $\times$  3 (type of activity; baseball, ballet, gender-neutral) ANOVAs were conducted, with repeated measures on the last factor on the 5-point scale responses (1 = "only girls," 5 = "only boys"). Because they did not differ, analysis of the two gender-neutral activities (books and tennis) were combined. The analysis revealed a significant main effect of type of activity,  $F(2, 248) = 652.77, p < .001, \eta^2 = .84$ . Follow-up comparisons showed that the children evaluated ballet,  $M = 1.57$ , as significantly more likely to be performed by girls than the neutral activities,  $M = 3.10, F(1, 248) = 473.69, p < .001, \eta^2 = .66$ , and also rated baseball,  $M = 4.42$ , as significantly more likely to be performed by boys than the neutral activities,  $F(1, 248) = 319.46, p < .001, \eta^2 = .56$ . There was also a Gender  $\times$  Activity interaction,  $F(4, 248) = 3.47, p < .01, \eta^2 = .05$ , which indicated that first graders,  $M = 1.30$ , saw ballet as more likely to be performed by girls than did fourth,  $M = 1.73$ , or seventh graders,  $M = 1.63$  and that first graders,  $M = 3.25$ , also rated the neutral activities (books and tennis) as more likely to

be performed by boys than did fourth,  $M = 3.00$ , or seventh graders,  $M = 2.97$ .

To investigate whether children held the appropriate beliefs about the race activities,  $2$  (gender of child)  $\times$   $3$  (grade)  $\times$   $3$  (type of activity; math, basketball, race-neutral) ANOVAs were conducted, with repeated measures on the last factor on the 5-point scale responses ( $1$  = "only Black children,"  $5$  = "only White children"). Because they did not differ, analysis of the two race-neutral activities (art and singing) were combined. The analysis revealed only a main effect of activity,  $F(2, 248) = 42.75$ ,  $p < .001$ ,  $\eta^2 = .26$ . (There were no significant findings for grade level or gender of the child.) Follow-up analyses showed that ratings for basketball,  $M = 2.52$ , were significantly lower than the neutral items,  $M = 3.19$ ,  $F(1, 248) = 47.25$ ,  $p < .001$ ,  $\eta^2 = .16$ , although the math rating,  $M = 3.43$ , was not significantly higher than the rating of the neutral items,  $F(1, 248) < 1.00$ . Taken together, then, these analyses indicate that children were aware of three out of the four stereotypes (only math was not significantly different from the neutral ratings).

### Straightforward Exclusion

It was expected that exclusion in the straightforward context would be seen as unwarranted by all children and that the inappropriateness would be justified by using moral reasons because the social-conventional aspect (some children would feel uncomfortable having the stereotyped child in the group) was not very strong. This hypothesis was tested by using  $2$  (gender of child)  $\times$   $3$  (grade)  $\times$   $2$  (group excluded: gender, race) ANOVAs with repeated measures on the last factor. There were no significant differences for these judgments across either gender of child or grade level. As expected, even though children possessed the appropriate knowledge about the activities, they did not use this knowledge to justify the exclusion as legitimate in the straightforward context. Rather, the vast majority of children,  $M = .96$ , judged that it was wrong for the peer groups to exclude a child from the activity across all contexts (see Table 2). In addition, children justified their decisions by using moral criteria. All children gave primarily moral,  $M = .92$ , rather than social-conventional,  $M = .06$ , justifications (see Table 2), and this tendency did not differ across either gender of child or grade level. There was also a main effect of type of exclusion,  $F(1, 124) = 5.48$ ,  $p < .05$ ,  $\eta^2 = .04$ , which indicates that children judged exclusion as more wrong in the race,  $M = .98$ , than the gender,  $M = .94$ , stories. This result appeared in every other analysis of both judgments and ratings, across each of the

**Table 2** Judgments, Justifications, and Ratings by Excluded Group and Grade: Straightforward Context

Excluded Group	Grade		
	1st	4th	7th
Gender exclusion			
Judgments	.94 (.17)	.94 (.19)	.93 (.20)
Ratings	4.86 (1.47)	4.24 (1.30)	4.17 (1.25)
Justifications			
Moral	.89 (.23)	.91 (.25)	.88 (.25)
Social-conventional	.09 (.20)	.08 (.24)	.11 (.25)
Race exclusion			
Judgments	.95 (.19)	.97 (.12)	1.00 (.13)
Ratings	4.94 (1.40)	4.49 (1.19)	4.97 (1.02)
Justifications			
Moral	.90 (.23)	.92 (.21)	.99 (.08)
Social-conventional	.05 (.19)	.04 (.13)	.01 (.08)

*Note:* Judgments are mean proportions of children stating "Not alright to exclude" (0 = alright; 1 = not alright). Ratings: 0 (not at all bad) to 6 (very, very bad) for excluding someone. Justifications are mean proportions of children using moral and social-conventional reasons for their judgment. Standard deviations are in parentheses.

three contexts. This finding is hard to interpret, however, because the activities from which the children were excluded also varied between gender and race groups. Therefore, it is not reported again.

The ratings of how bad it would be to exclude a child reflected a pattern similar to the judgment data. All children judged such exclusion to be wrong,  $M = 4.59$  out of 6.00 overall, and there were no grade differences on these ratings. There was a main effect, however, of gender of child,  $F(1, 124) = 9.28$ ,  $p < .001$ ,  $\eta^2 = .07$ , which showed that girls,  $M = 4.89$ , rated exclusion as more bad than boys,  $M = 4.29$ ,  $p < .001$ . There was also an unexpected grade  $\times$  type of exclusion interaction,  $F(2, 124) = 7.76$ ,  $p < .001$ ,  $\eta^2 = .11$ . As revealed in Table 2, this interaction showed that children's negative ratings of exclusion in the gender scenarios decreased with age,  $Ms = 4.86, 4.24, 4.17$  for first, fourth, and seventh grades,  $p < .001$ , whereas there was no corresponding pattern for the ratings about race.

In sum, strong support was found for the expectation that all children would see straightforward exclusion as wrong and that they would justify these decisions with moral reasoning. Even the youngest children were quite aware of the inappropriateness of exclusion, despite the fact that they judged the activities as stereotypically inappropriate for the child. Also, as expected, girls saw this exclusion as more wrong overall than did boys. A separate analysis was conducted for the two gender-related stories but no evidence was found for ingroup favoritism; that is,

girls and boys did not judge that it was more all right for a group to exclude someone from the opposite sex than from one's own sex.

### Multifaceted Contexts

Children's choices in the multifaceted contexts involved choosing which of two children should join the club when only one place was left and when there were competing reasons for the decision. Thus, in these contexts children needed to weigh both moral and social conventional considerations to make their decisions. In the equal qualifications context, both children were said to have equal skills at the activity. In the unequal qualifications context, the child who fit the activity was also said to be better at it, thus making salient the potential cost to effective functioning of the group if the nonstereotypical child was included. The hypotheses were tested by using 2 (gender of child)  $\times$  3 (grade)  $\times$  2 (group excluded; gender, race)  $\times$  2 (context: equal qualifications, unequal qualifications) ANOVAs, with repeated measures on the last two factors. Because it was found that the patterns of the dependent variables did not differ for race versus gender exclusion, this factor was collapsed across for the remaining analyses. Thus, 2 (gender of child)  $\times$  3 (grade level)  $\times$  2 (context: equal qualifications, unequal qualifications) ANOVAs with repeated measures were conducted on the last factor.

**Choices.** Across all of the multifaceted contexts (equal and unequal qualifications), the majority of the children,  $M = .60$ , favored including the nonstereotypical child (that is, the boy for ballet, the girl for baseball cards, the Black child for math, and the White child for basketball) as shown in Table 3. Demonstrating, however, that the manipulation of the salience of social conventional factors was successful, this tendency was significantly greater in the equal,  $M = .74$ , than in the unequal qualifications context,  $M = .52$ ,  $F(1, 124) = 61.33$ ,  $p < .001$ ,  $\eta^2 = .33$ . There was also a main effect of gender of child,  $F(1, 124) = 6.15$ ,  $p < .001$ ,  $\eta^2 = .05$ , which indicates that, across both contexts, girls,  $M = .67$ , were more likely to choose the nonstereotypical child than were boys,  $M = .52$ ,  $p < .001$ .

More importantly, the expected Grade Level  $\times$  Context (equal or unequal qualifications) interaction emerged,  $F(2, 124) = 6.15$ ,  $p < .01$ ,  $\eta^2 = .09$ , on the choices. Suggesting that they were not sensitive to the potential negative effects of including an unskilled child in the group, the fourth-grade children did not significantly differentiate between the unequal,  $M = .77$ , and equal,  $M = .74$ , contexts,  $p > 1.00$ . The seventh-grade children, however, did rate exclusion less neg-

**Table 3 Choices and Ratings by Children Who Chose the Child Who Did Not Fit the Stereotype in the Multifaceted Contexts**

Excluded Group by Context	Grade		
	1st	4th	7th
Gender exclusion			
Choices			
Equal qualifications	.71 (.39)	.67 (.40)	.60 (.46)
Unequal qualifications	.41 (.46)	.55 (.46)	.22 (.40)
Ratings			
Equal qualifications	2.88 (1.90)	2.49 (1.64)	2.94 (1.40)
Unequal qualifications	2.83 (1.90)	2.35 (1.65)	2.05 (1.60)
Racial exclusion			
Choices			
Equal qualifications	.79 (.32)	.88 (.28)	.77 (.37)
Unequal qualifications	.51 (.46)	.74 (.40)	.29 (.43)
Ratings			
Equal qualifications	3.49 (1.83)	3.79 (1.58)	3.51 (1.75)
Unequal qualifications	3.40 (1.91)	3.56 (1.73)	2.44 (1.70)

*Note:* Judgments are mean proportions of children who chose the child who did not fit the stereotype (0 = child who fit the stereotype; 1 = child who did not fit the stereotype). Ratings: 0 (not bad) to 6 (very, very bad) for not picking the child who did not fit the stereotype. Standard deviations are in parentheses.

atively in the unequal qualifications context in which the unskilled child might have influenced effective group functioning,  $F(1, 42) = 33.1$ ,  $p < .001$ ,  $\eta^2 = .44$ ,  $Ms = .69$  and  $.29$ , respectively. Unexpectedly, the first-grade children were also more likely to choose the nonstereotypical child in the equal,  $M = .75$ , than the unequal,  $M = .50$ , contexts,  $F(1, 38) = 10.35$ ,  $p < .01$ ,  $\eta^2 = .21$ . There were no main effects for either grade or gender on this measure. Moreover, there was no evidence for an ingroup bias (girls did not choose girls in girl-typed contexts more than did boys; nor did boys choose boys in boy-typed contexts more than did girls).

**Justifications.** The justification measure provided the most important test of the expectation that social-conventional reasoning about group functioning varied across age. Overall, it was found that children used more moral,  $M = .67$ , than social-conventional,  $M = .29$ , justifications,  $F(1, 124) = 46.02$ ,  $p < .001$ ,  $\eta^2 = .27$ ; however, children used more moral than social-conventional justifications in the equal,  $Ms = .67$  and  $.35$ , respectively, than in the unequal contexts,  $Ms = .50$  and  $.41$ ,  $F(1, 124) = 22.78$ ,  $p < .001$ ,  $\eta^2 = .26$ . Furthermore, although the proportion of social conventional versus moral justifications increased across grade,  $F(2, 124) = 5.92$ ,  $p < .05$ ,  $\eta^2 = .09$  in both contexts, this tendency was significantly greater in the unequal than the equal context contexts; the expected Grade  $\times$  Context  $\times$  Type of Justification interaction was significant,  $F(2, 124) = 4.08$ ,  $p < .05$ ,  $\eta^2 = .06$ . The

**Table 4** Children's Justifications for Their Choices in the Multi-faceted Contexts

Excluded Group by Context	Grade		
	1st	4th	7th
Gender exclusion			
Equal qualifications			
Moral	.64 (.41)	.59 (.41)	.51 (.40)
SC	.31 (.31)	.39 (.41)	.49 (.40)
Unequal qualifications			
Moral	.55 (.40)	.54 (.43)	.37 (.37)
SC	.30 (.36)	.41 (.41)	.62 (.38)
Racial exclusion			
Equal qualifications			
Moral	.69 (.37)	.81 (.34)	.77 (.37)
SC	.19 (.32)	.13 (.30)	.22 (.37)
Unequal qualifications			
Moral	.55 (.40)	.65 (.41)	.35 (.43)
SC	.26 (.33)	.27 (.39)	.64 (.44)

Note: Justifications are proportions of children using moral and social-conventional reasons for their choices. Moral = moral justifications; SC = social-conventional justifications. Standard deviations are in parentheses.

results of this analysis are shown in Table 4. This interaction is also charted in Figure 1, which represents a difference score between moral and social conventional reasoning at each age level and for each context such that higher numbers indicate a greater proportion of moral (versus social-conventional) reasoning. The first and fourth graders were not as sensitive to context as were the seventh graders and gave more moral than social conventional justifications for both the equal and unequal contexts. The seventh graders, on the other

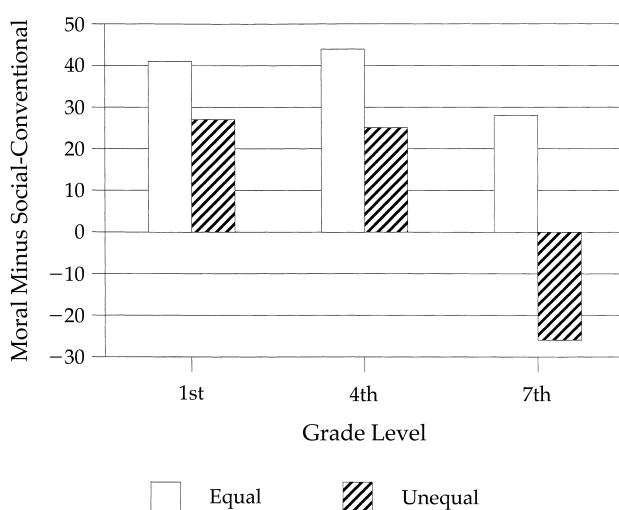
hand, viewed the equal context in moral terms and the unequal context in social-conventional terms.

To determine the nature of the justifications, the use of the three subcategories of social-conventional justifications was examined. In the unequal contexts, the vast majority of justifications,  $M = .83$ , were about group functioning (with the remaining justifications being about conventions,  $M = .08$ , and group identity,  $M = .09$ ). In fact, for the seventh-grade children, the vast majority,  $M = .92$ , of the social-conventional justifications in the unequal context were about group functioning. When the nonstereotypical child threatened group functioning, the older children were less willing to choose this child to participate. In the equal qualifications contexts, however, in which "all things were equal," most participants used moral justifications (and chose someone who did not fit the stereotype of the activity). Of those participants who used social-conventional justifications in the equal qualifications contexts, only a minority used ones based on group functioning,  $M = .11$ ; most participants gave justifications based on conventions,  $M = .51$ , and group identity,  $M = .38$ .

**Ratings.** Across all of the multifaceted contexts, the mean rating (how bad would it be to exclude the nonstereotypical child?), on a 0 (not at all bad) to 6 (very, very bad) scale, was  $M = 2.98$ . These data are depicted in Table 3. Excluding the nonstereotypical child, however, was seen as less bad in the unequal,  $M = 3.18$ , than in the equal qualifications context,  $M = 2.77$ ,  $F(1, 124) = 23.24$ ,  $p < .001$ ,  $\eta^2 = .16$ . Again, the expected Grade  $\times$  Context interaction emerged,  $F(2, 124) = 11.07$ ,  $p < .01$ ,  $\eta^2 = .15$ , on the ratings. Neither the first- nor the fourth-grade children significantly differentiated between the unequal and equal contexts,  $M_s = 3.18$  and  $3.11$  for the first graders and  $3.14$  and  $2.95$  for the fourth graders, both  $p > .16$ . The seventh-grade children, however, did differentiate between the two contexts,  $F(1, 42) = 40.19$ ,  $p < .001$ ,  $\eta^2 = .49$ ,  $M_s = 3.22$  and  $2.24$ , respectively, rating it as more bad to exclude the stereotypical child in the equal than the unequal contexts. Again there were no main effects for either grade or gender on this measure.

#### Relation between Judgments and Justifications

The analyses are based on the theory that children are more likely to allow exclusion on the basis of gender or race when they perceive the issue to be one involving social-conventional rather than moral considerations. To examine the hypothesis that decisions to exclude individuals from activities would be justified primarily by using social-conventional justifications, whereas decisions of inclusion would be justified



**Figure 1** The proportion of moral minus the proportion of social-conventional justifications given by children at each grade for equal and unequal contexts.

primarily by using moral reasoning, the proportion of moral and social-conventional justifications given by children who picked the stereotypical or nonstereotypical child were calculated (the straightforward context was not included in this analysis because the vast majority of all participants stated that it was wrong to exclude and gave moral justifications).

Confirming the expected relation between judgment and justification, across both equal and unequal contexts, children who picked the stereotypical child used a greater proportion of moral,  $M = .87$ , than social-conventional,  $M = .17$ , reasoning,  $p < .001$ . Children who picked the nonstereotypical child, however, used a greater proportion of social-conventional,  $M = .85$ , than moral,  $M = .04$ , reasoning,  $p < .001$ .

## DISCUSSION

Overwhelmingly, children and adolescents rejected straightforward exclusion on the basis of gender or race, even in contexts in which gender or racial stereotypic expectations could provide a basis for condoning exclusion (such as excluding a boy from ballet because ballet is an activity primarily performed by girls). The reasons given for rejecting exclusion were primarily fairness and rights, equal treatment, and equal access. Thus, children interpreted straightforward exclusion, even in stereotypic contexts, as wrong from a moral viewpoint, which supports other findings regarding children's judgments of straightforward violations of rights and equality (Helwig, 1995, 1997; Killen, 1991; Smetana, 1995; Tisak, 1995).

Also supporting predictions derived from social-cognitive developmental theories, in multifaceted situations in which participants were asked to pick one of two children to join a peer group club, children weighed a variety of issues to make their decisions of whom to pick. The children were influenced by considerations of the qualifications of the children (how good they were at the club's activity) as well as issues of fairness and equal opportunity. When a choice had to be made between two equally qualified children, most of the participants stated that the child who did not fit the stereotype should be picked and gave reasons based on equal treatment and equal access. Justifications such as "Boys don't get a chance to take ballet" and "A girl could learn how to play baseball cards and then she could teach other girls how to play" were used in the gender contexts. Similarly, for the race stories, children who picked the child who did not fit the stereotype used reasons such as "It's good for Black and White children to play together and learn about each other so they'll get along well and not be prejudiced."

These results are different than those from previous research that used a similar methodology in preschool-aged children (Theimer et al., 2001) and in which a majority of young children chose a child who fit the stereotype (e.g., a girl for ballet) in contexts in which the qualifications were equal. Perhaps older children are more sensitive to considerations involving prior history of opportunity, whereas young children may rely more upon their beliefs about group-appropriate activities to make a decision about whom to pick (influencing them, for instance, to pick a girl instead of a boy for ballet). Supporting this interpretation was the finding that first graders were more likely to identify ballet as a girls' activity than were fourth and seventh graders. However, the contexts used in our study and the preschool study just cited were somewhat different, so further research will be needed to substantiate this interpretation.

Although the activity knowledge measure confirmed that the children were well aware of differences in the appropriateness of the activities across racial and gender groups, it was found that only a very few children (less than 8% overall) used stereotypes to justify their answers (and these justifications were found only in the equal opportunity gender contexts). These few responses included statements such as "It's okay to pick the girl for ballet because boys aren't good at ballet." In short, moral and other types of social conventional reasons appear to have been more important to these children than stereotypes about appropriate peer group activities as judged by their use of justifications. Unexpectedly, first-grade children were more likely to pick the nonstereotypic child in the equal qualifications context than were older children even though they were also more likely to identify ballet as a "girls' activity" in the group activity knowledge assessment. Thus, the younger children did not focus on the group-functioning dimension of the multifaceted contexts to the extent that older children did despite their stronger associations of gender with the types of group activities being discussed. This could be a lack of coordination of different aspects of social knowledge or it could pertain to their concepts of group functioning. Further research is needed to fully interpret this pattern of results.

The seventh graders were most sensitive to context in that they differentiated between the equal and unequal qualifications contexts. For the dimension of a child's qualifications to join a group, the older children seemed to be more aware of the potential cost to group functioning of including a child who was not qualified than were the younger children. Although issues of fairness prevailed overall, older children qualified these judgments in cases where there was a

potential threat to group functioning. In the unequal qualifications contexts, the seventh graders were more likely to choose the more qualified child to join the club and to use social-conventional reasons (particularly those relating to group functioning) to justify these choices. An exception to this pattern occurred on the choice measures, in which the first graders also differentiated between the equal opportunity and the unequal opportunity contexts (although it did not occur on ratings or justifications). Although there is no clear explanation for this pattern, it is similar to previously documented u-shaped developmental curves for children's social-conventional knowledge about gender roles (see Carter & Patterson, 1982; Stoddart & Turiel, 1985). Turiel's (1983) findings on children's social-conventional knowledge has revealed an affirmation–negation fluctuation, in which children affirm and then negate conventions throughout development so that curvilinear patterns result. No curvilinear pattern was found for our group activity knowledge, however, and so more work on this finding is warranted.

Prior research by Damon (1977, 1983) found that children used justifications based on individual merit (such as "He deserves to be in the club because he is good at it") and that older children take merit and effort into consideration more than younger children when deciding how to divide resources. In this study, children who picked a child who was better at the activity than another child took the group-functioning considerations into account more than they did the individual merit of the child (less than 6% of the justifications were individual merit). Apparently, group functioning (making the club better) was a more salient feature of this type of decision making than was individual merit. Reasoning about individual merit may be more prominent for decisions involving the distribution of resources or rewards than for decisions about inclusion and exclusion. A child's qualifications, however, is only one consideration that bears on children's judgments about group functioning and social group processes. Further research is needed to examine the many other relevant dimensions that bear on judgments about exclusion. In addition, it would be important to examine children's developmental knowledge about stereotypes in a more detailed way than was done in the present study to analyze direct relationships between children's developmental knowledge about stereotypes and their judgments about exclusion.

In general, it was found that children viewed exclusion by race to be more inappropriate than exclusion by gender. Although this could indicate a greater concern with racial exclusion, the activities that were

used in the race and gender contexts were not the same (math and basketball for the race activities, ballet and baseball for the gender activities), so it is difficult to draw firm conclusions about this. Other than this, however, little evidence was found that children's judgments about inclusion and exclusion were influenced either by the race or gender of the child being excluded or by the relation between one's own category membership and that of the child being excluded (girls did not rate girl-excluded contexts as more wrong than boy-excluded contexts nor did boys rate boy-excluded contexts as more wrong than girl-excluded contexts). Thus, the children did not see it as differentially wrong to exclude White children versus Black children from activities or to exclude girls versus boys from activities. In short, children did not display ingroup favoritism (consistent with previous research on the decline of prejudice during the middle-school years; see Aboud, 1988). Further research is needed to determine whether children show a concern with the potential for "affirmative action" type decisions, in which groups that have traditionally been excluded might be given more access.

Overall, and as expected, girls were more concerned with fairness and equal access than were boys. In the straightforward scenarios, they rated exclusion as more wrong, and in the multifaceted contexts they were more likely to want to include the child who was nonstereotypical for the activity (that is, the child who had less experience). These findings support previous research indicating that girls rate helping and caring as more obligatory than do boys in certain contexts (Killen & Turiel, 1998; Wentzel & Erdley, 1993; Zarbatany et al., 1996). Girls may be more sensitive to exclusion than boys on the basis of their own experience of being excluded from gender-specific activities such as sports. Children and adolescents who are members of groups that have been historically excluded from participating in group activities may be more sensitive to issues of inclusion. This hypothesis, however, remains to be tested in future work.

The sample of children in this study were from a mixed-ethnicity, middle-class, suburban city outside of Washington, DC. It is feasible that this environment would facilitate an awareness of the wrongfulness of holding and using prejudice and stereotypes. To test this hypothesis, further research in other communities, and in more homogeneous regions, needs to be conducted to fully understand how children evaluate group inclusion and exclusion, particularly on the basis of gender or race. Preliminary analyses with a small sample of African American children indicated that African American children were more

likely to pick a child (boy or girl; Black or White child) who did not fit the stereotype than were the children reported in this paper. This also suggests that children who experience exclusion and who may be targets of discrimination may be more sensitive to decision making involving inclusion and exclusion, consistent with the finding that girls were more concerned with fairness and equal access than were boys. To systematically test this hypothesis, research has to be conducted with children who are from different ethnic backgrounds and who have interacted with others who differ from members of their own group. To adequately examine the role of experience, however, comprehensive measures of personal experience of exclusion, discrimination, and contact with others have to be developed and employed. In the authors' view, inferences about children's experience with exclusion cannot be made solely on the basis of their gender or racial group membership but, rather, should be made as a function of self-perception and of actual recorded experience. Further, a child's degree of experience of interacting with others from different social groups may or may not contribute to a sensitivity about exclusion. Again, detailed measures of this type of experience have to be documented and analyzed in relation to children's social judgments and evaluations of inclusion and exclusion.

In sum, the children in this study judged that it was wrong to exclude someone from a peer group activity on the basis of gender or race. In multifaceted contexts, however, other considerations, such as group functioning, entered into children's judgments, and particularly so with age. In this sense, it appears that judgments about the appropriateness of exclusion on the basis of gender or race in peer group activities may not be based so much upon direct prejudice or stereotypes about those groups or group members but rather more indirectly in terms of the perceived costs of excluding individuals from group activities. Other contexts, such as the family, school, and workplace, may be very different and require further investigation. Future work involving a wider range of contexts, observational data, and reports from teachers and parents is needed to provide more insight into why social group functioning becomes important with age, and how it affects reasoning about intergroup relationships.

## ACKNOWLEDGMENTS

This research was supported by grant #9729739 from the National Science Foundation awarded to both authors. A project grant from the University of Mary-

land, College Park, Graduate Research Board, awarded to the first author, provided additional funding. The authors thank George Bregman, Elizabeth Hakola, Robert Haxter, Andrew Herst, Stacey Horn, B. Sefton Price, and Christine Theimer Schuette for their assistance with data collection and data transcription. In addition, Stacey Horn and Christine Theimer Schuette conducted interrater reliability. The authors extend their gratitude to Elliot Turiel for feedback on the manuscript and to the anonymous reviewers. They also thank Joseph Hawkins and the Superintendent's Office at the Montgomery County Public Schools for allowing us to conduct this study in the public schools. Much appreciation is extended to the teachers, children, and parents for their participation.

## ADDRESSES AND AFFILIATIONS

Corresponding author: Melanie Killen, Department of Human Development, 3304 Benjamin Building, University of Maryland, College Park, MD 20742-1131; email: mk141@umail.umd.edu. Charles Stangor is also at the University of Maryland.

## REFERENCES

- Aoud, F. E. (1988). *Children and prejudice*. New York: Basil Blackwell.
- Aoud, F. E. (1992). Conflict and group relations. In C. U. Shantz & W. W. Hartup (Eds.), *Conflict in child and adolescent development* (pp. 356-379). Cambridge, U.K.: Cambridge University Press.
- Asher, S., & Coie, J. (1990). *Peer rejection in childhood*. Cambridge, U.K.: Cambridge University Press.
- Bennett, M., Barrett, M., Lyons, E., & Sani, F. (1998). Children's subjective identification with the group and in-group favoritism. *Developmental Psychology, 34*, 902-909.
- Brewer, M. B. (1979). In-group bias in the minimal intergroup situation: A cognitive-motivational hypothesis. *Psychological Bulletin, 86*, 307-324.
- Brown, B. (1989). The role of peer groups in adolescent's adjustment to secondary school. In T. Berndt & G. Ladd (Eds.), *Peer relationships in child development* (pp. 188-215). New York: Wiley.
- Carter, D. B., & Patterson, C. J. (1982). Sex roles as social conventions: The development of children's conceptions of sex-role stereotypes. *Developmental Psychology, 18*, 812-824.
- Damon, W. (1977). *The social world of the child*. San Francisco: Jossey-Bass.
- Damon, W. (1983). *Social and personality development*. New York: Norton.
- Eisenberg, N., & Miller, P. (1987). The relation of empathy to prosocial and related behaviors. *Psychological Bulletin, 101*, 91-119.

Harter, S. (1998). The development of self-representations. In N. Eisenberg (Ed.), W. Damon (Series Ed.), *Handbook of child psychology: Vol. 3. Social, emotional, and personality development* (5th ed., pp. 553–617). New York: Wiley.

Helwig, C. C. (1995). Adolescents' and young adults' conceptions of civil liberties: Freedom of speech and religion. *Child Development*, 66, 152–166.

Helwig, C. C. (1997). The role of agent and social context in judgments of freedom of speech and religion. *Child Development*, 68, 484–495.

Killen, M. (1991). Social and moral development in early childhood. In W. M. Kurtines & J. L. Gewirtz (Eds.), *Handbook of moral behavior and development* (Vol. 2, pp. 115–138). Hillsdale, NJ: Erlbaum.

Killen, M., & Turiel, E. (1998). Adolescents' and young adults' evaluations of helping and sacrificing for others. *Journal of Research on Adolescence*, 8, 355–375.

Liben, L. S., & Signorella, M. L. (1993). Gender-schematic processing in children: The role of initial interpretations of stimuli. *Developmental Psychology*, 29, 141–149.

Maccoby, E. (1988). Gender as a social category. *Developmental Psychology*, 24, 765–775.

Mackie, D. M., Hamilton, D. L., Susskind, J., & Rosselli, F. (1996). Social psychological foundations of stereotype formation. In C. Macrae, C. Stangor, & M. Hewstone (Eds.), *Stereotypes and stereotyping* (pp. 41–77). New York: Guilford Press.

Macrae, C., Stangor, C., & Hewstone, M. (Eds.). (1996). *Stereotypes and stereotyping*. New York: Guilford Press.

Short, G. (1993). Sex-typed behavior in the primary school: The significance of contrasting explanations. *Educational Research*, 35, 77–87.

Smetana, J. G. (1995). Morality in context: Abstractions, ambiguities, and applications. In R. Vasta (Ed.), *Annals of child development* (Vol. 10, pp. 83–130). London: Jessica Kingsley.

Stangor, C., & Ruble, D. N. (1989). Differential influences of gender schemata and gender constancy on children's information processing behavior. *Social Cognition*, 7, 353–372.

Stoddart, T., & Turiel, E. (1985). Children's concepts of cross-gender activities. *Child Development*, 56, 1241–1252.

Tajfel, H., Billig, M., Bundy, R. P., & Flament, C. (1971). Social categorization and intergroup behavior. *European Journal of Social Psychology*, 1, 149–177.

Theimer, C. E., Killen, M., & Stangor, C. (2001). Preschool children's evaluations of exclusion in gender-stereotypic contexts. *Developmental Psychology*, 37, 1–10.

Tisak, M. (1995). Domains of social reasoning and beyond. In R. Vasta (Ed.), *Annals of child development* (Vol. 11, pp. 95–130). London: Jessica Kingsley.

Turiel, E. (1978). The development of concepts of social structure: Social convention. In J. Glick & K. A. Clarke-Stewart (Eds.), *The development of social understanding* (pp. 25–107). New York: Gardner Press.

Turiel, E. (1983). *The development of social knowledge: Morality and convention*. Cambridge, U.K.: Cambridge University Press.

Turiel, E. (1998). The development of morality. In N. Eisenberg (Ed.), W. Damon (Series Ed.), *Handbook of child psychology: Vol. 3. Social, emotional, and personality development* (5th ed., pp. 863–932). New York: Wiley.

Turiel, E., Killen, M., & Helwig, C. C. (1987). Morality: Its structure, functions, and vagaries. In J. Kagan & S. Lamb (Eds.), *The emergence of morality in young children* (pp. 155–244). Chicago: University of Chicago Press.

Van Avermaet, E., & McClintock, L. G. (1988). Intergroup fairness and bias in children. *European Journal of Social Psychology*, 18, 407–427.

Winer, B. J. (1971). *Statistical principles in experimental design*. New York: McGraw-Hill.

Wentzel, K., & Erdley, C. (1993). Strategies for making friends: Relations to social behavior and peer acceptance in early adolescence. *Developmental Psychology*, 29, 819–826.

Yee, M. D., & Brown, R. J. (1992). Self evaluations and intergroup attitudes in children aged three to nine. *Child Development*, 63, 619–629.

Zahn-Waxler, C., Cole, P. M., Welsh, J. D., & Fox, N. (1995). Psychophysiological correlates of empathy and prosocial behavior in preschool children with behavior problems. *Development and Psychopathology*, 1, 27–48.

Zarbatany, L., Van Brunschot, M., Meadows, K., & Pepper, S. (1996). Effects of friendship and gender on peer group entry. *Child Development*, 67, 2287–2300.

This document is a scanned copy of a printed document. No warranty is given about the accuracy of the copy.  
Users should refer to the original published version of the material.