European American children’s and adolescents’ evaluations of interracial exclusion

Melanie Killen, Megan Clark Kelly, Cameron Richardson, David Crystal, and Martin Ruck

Abstract
No research, to date, has investigated the role of ethnic school composition (and intergroup contact) on European American youth’s use of stereotypes to explain interracial discomfort in the context of peer exclusion. In this study, European American fourth-, seventh- and 10th-grade students (N = 414), attending low and high ethnically-diverse public schools (with low and high self-reports of cross-race/ethnic friendships respectively) evaluated three contexts of interracial exclusion (at lunch time, at a school dance, and at a sleepover). In addition to age and context effects, participants enrolled in high-diversity schools were less likely to use stereotypes to explain racial discomfort, more likely to view racial exclusion as wrong, and more likely to estimate that racial exclusion occurs, than were participants enrolled in low-diversity schools. These findings have implications for the role of social experience on racial attitudes and judgments about exclusion.

Keywords
child development, exclusion, intergroup contact, morality, stereotypes

Paper received 13 November 2008; revised version accepted 17 July 2009.

Extensive research has been conducted with adult populations on judgments about intergroup inclusion and exclusion (Abrams, Hogg, & Marques, 2005), and recently developmental psychologists have investigated children’s and adolescents’ evaluations of inclusion and exclusion (Abrams, Rutland, Pelletier, & Ferrell, 2009; Horn, 2003; Killen, Henning, Kelly, Crystal, & Ruck, 2007; Nesdale, 2004; Rutland, 2004). An intergroup approach to examining exclusion in childhood differs from the large body of research on peer rejection in the developmental psychology literature. The latter has focused predominantly on personality traits that lead to rejection (such as shyness or aggression) rather than on dimensions that reflect prejudice, bias, or stereotyping (see Killen, Sinno, & Margie, 2007). For example, the developmental peer relations...
literature examines how individual differences in personality traits lead to negative social interactions, with an emphasis on how to intervene to improve children's social skills and reduce the negative consequences of exclusion (see Rubin, Bukowski, & Parker, 2006).

By contrast, research on intergroup attitudes and relations, drawing on social psychology, has provided a new lens for understanding the developmental source of childhood exclusion based on prejudice and stereotyping, as well as for determining the types of intervention that may improve peer relationships (see Killen, Rutland, & Jampol, 2008). One area of this research has focused on social reasoning about exclusion. The research findings, to date, have indicated that children and adolescents often use moral reasoning (such as fairness) when evaluating gender and racial exclusion in straightforward situations, but rely on social conventional and personal reasoning to justify exclusion in ambiguous or complex situations. This supports social psychological findings with adults reflecting that prejudice is more likely to be activated in ambiguous rather than in explicit social contexts (Dovidio & Gaertner, 2004). Social conventional reasoning in childhood and adolescence typically refers to the use of justifications that reflect group functioning (“The group won’t work well with someone different in it”); personal reasoning refers to individual choice (“It’s my choice”). This taxonomy of social reasoning stems from social domain theory (Turiel, 2002) and has been applied to the topic of intergroup exclusion in childhood with a range of findings detailing the ways that judgments vary with the target and context (Abrams et al., 2009; Killen, 2007).

What has not been examined, however, is the extent to which children and adolescents use stereotypes to justify exclusion, and how this might vary with the social context, as well as change with age. Further, while extensive research in the adult literature has been devoted to the ways that contact with members of outgroups reduces prejudice (Hewstone, Rubin, & Willis, 2002; Pettigrew & Tropp, 2005), much less research has examined the role of intergroup contact on the use of stereotypes in childhood (for a review, see Tropp & Prenovost, 2008). Intergroup contact theory postulates that intergroup contact, alone, is not enough to reduce prejudice. Certain conditions must be met. The condition of contact that has been demonstrated to be the most robust factor to reduce prejudice is cross-group (cross-race) friendship (Pettigrew & Tropp, 2005); however, the majority of intergroup contact research has been conducted with adults (see Tropp & Prenovost, 2008, for an exception).

**School diversity and social context**

Studies with children that have been conducted with intergroup contact have shown that enrollment in heterogeneous schools is associated with positive effects such as less racial prejudice as measured by trait-assignment tasks (Rutland, Cameron, Bennett, & Ferrell, 2005), greater focus on shared interests than race when making judgments about similarity between peers (McGlothlin & Killen, 2005), and greater sense of wrongfulness regarding interracial exclusion (Crystal, Killen, & Ruck, 2008). Additionally, researchers have demonstrated social developmental benefits for adolescents attending ethnically-diverse schools, such as increased perceptions of school safety and decreased perceptions of vulnerability (Hanish & Guerra, 2000; Juvonen, Nishina, & Graham, 2006).

Surprisingly, no research has investigated whether diverse school experiences are significantly related to the use of stereotypes in explaining interracial peer encounters. Investigating the role of school diversity and context on use of stereotypes has implications for intergroup contact research as well as for school desegregation research, which is concerned with understanding how school composition, in terms of ethnicity, impacts on children’s race-related attitudes (Frankenberg & Orfield, 2007; Killen, Henning, et al., 2007; Orfield, 2001; Schofield, 1995). Endorsement of stereotypes has been widely associated with prejudice, discrimination, and
racial bias in both children and adults (e.g., Corenblum & Stephan, 2002; McKown, 2004; Steele, Choi, & Ambady, 2004). The school desegregation literature has focused on whether school diversity improves test scores, reduces delinquency, and reduces prejudicial attitudes as defined by trait-assignment tests (Frankenberg & Orfield, 2007). Inclusion of stereotype-related measures is consistent with the viewpoint of Frankenberg and Orfield (2007), who have asserted that developmental social cognitive research is making a significant contribution to general knowledge about children’s intergroup attitudes (Frankenberg & Orfield, 2007).

The present study

Prior research among youth has found that situations describing interracial exclusion are potent stimuli for eliciting various forms of intergroup judgments and attitudes (Abrams et al., 2009; Killen, 2007; Killen, Lee-Kim, McGlothlin, & Stangor, 2002). The focus of the present study, then, was on how majority youth, specifically European American (“White”) children and adolescents, evaluated interracial exclusion scenarios, and the extent to which they invoked stereotypes to explain why some individuals might experience discomfort in interracial social interactions.

Previous studies have shown that the social context makes a difference in children’s and adolescents’ social evaluations of exclusion (Abrams & Rutland, 2008; Killen, Henning, et al., 2007). Judgments about exclusion vary as a function of the target of exclusion (gender, race) as well as whether exclusion occurs in school or home settings. With the exception of a few studies, what has not been done is to vary the intimacy of the context. Drawing on an interview method established in a previous study (Crystal et al., 2008), the contexts in the present investigation included two peer encounters related to school (eating together at lunch time, and attending a school dance) and one peer encounter related to the home (bringing a friend home for a sleepover party). These contexts reflected familiar interpersonal exchanges for children and adolescents, including the involvement of parental expectations (sleepover party). Given that racial integration remains fairly minimal in intimate contexts in the adolescent world with respect to cross-race friendship interactions such as a sleepover party or dating (Edmonds & Killen, 2009; Kennedy, 2003), and that cross-race relationships remain infrequent in the US (Page-Gould, Mendoza-Denton, & Tropp, 2008), we expected these contexts to elicit strong race-related responses from participants.

In these contexts, a majority group member (European American) excluded a minority group member (African American) and a mixture of factors were potential reasons for exclusion, among them, race, lack of shared interests (in the lunch encounter), anticipated parental unfamiliarity (in the sleepover encounter), and school rivalry (in the school dance context). The latter three reasons, lack of shared interests, parental unfamiliarity, and school rivalry, were referred to as “non race-based” reasons in contrast to race-based reasons (such as skin color) for exclusion in interracial peer encounters.

Non race-based reasons for exclusion (a lack of shared interests, parental unfamiliarity, and school rivalry) have been viewed by children and adolescents as legitimate reasons for peer rejection (Nucci, 2001; Smetana, 2005). Yet, in the context of interracial interactions, these forms of reasoning take on potentially different meaning than in same-race encounters. That is, when a White student does not want to have lunch with a Black student because “they have different sports interests” or when a White student does not invite a Black friend over to his/her house for a sleepover because “my parents may be uncomfortable” these “non race-based” reasons may be evaluated differently by individuals who have daily interracial interactions than by those who do not.

One aim of the study, then, was to determine whether reasons other than race were evaluated differently as a function of intergroup contact and age. Further, school identity and concerns about group functioning become quite important during middle and high school, and have been
shown to influence exclusion decisions, again, in situations in which race is not explicitly mentioned or included in the measurements (Horn, 2006; Nesdale, 2004). Thus, exclusion reasons based on group functioning were also included for evaluation.

The novel dimensions of the present project were to examine how majority participants’ evaluations of exclusion and use of stereotypes varied as a function of ethnic school composition (and intergroup contact), and whether there exist age-related changes in the way that school diversity impacts on participants’ judgments regarding interracial exclusion. We administered a measure to determine whether participants’ cross-race friendships provided more explicit information about contact with the outgroup than would be reflected by only using school diversity as the relevant variable (Crystal et al., 2008).

**Hypotheses**

In the present study, we expected that participants enrolled in ethnically-diverse schools (with more cross-race friendships) would be less likely to endorse or affirm stereotypes to explain why exclusion occurs in interracial contexts than would participants enrolled in schools without diversity (fewer cross-race friendships). This was expected on the basis of previous studies that have indicated that friendships with members of outgroups reduce prejudice (Cameron, Rutland, Brown, & Douch, 2006; Rutland et al., 2005; Tropp & Prenovost, 2008). We extended this finding by hypothesizing that positive contact would also reduce the endorsement of explicit stereotypes to explain racial discomfort and justify exclusion. For the same reasons that positive contact with members of outgroups reduces prejudice (through empathy and perspective-taking), we expected that positive outgroup contact would reduce the affirmation of stereotypes.

At the same time, contact with outgroup members may increase the recognition that stereotypes exist, and the awareness of the contexts that contribute to the use of stereotypes, such as a history of segregation. This contributes to our hypothesis that students enrolled in high-diversity schools may have a higher rate of recognition of stereotypes, and awareness of the social and historical contexts that explain the use of stereotypes, than students in low-diversity schools.

Additionally, we predicted age-related decreases in the endorsement of stereotypes in high-diversity school contexts due to the social cognitive awareness of what makes prejudice wrong, as an outcome of having cross-race friendships (Tropp & Prenovost, 2008). Further, parents frequently communicate their desire for same-race relationships in adolescence (Edmonds & Killen, 2009), suggesting that stereotypes might not decline with age for students at both high- and low-diversity schools for contexts such as a sleepover where parental discomfort with cross-race relationships may be a salient consideration for students.

We also predicted that the majority of all participants would view race-based exclusion as wrong, given prior research that has demonstrated such findings (Killen et al., 2002). Additionally, we expected that participants in high-diversity schools would judge non-race based reasons—such as lack of shared interests or parental unfamiliarity—for interracial exclusion to be more wrong than would participants in low-diversity schools. This hypothesis was based on previous findings in which minority students, who had higher levels of intergroup contact than did majority students, viewed non-race based reasons as more wrong than did majority students (Crystal et al., 2008; Killen, Henning, et al., 2007). Following this rationale, we expected that majority students with higher levels of intergroup contact, relative to those with lower levels of such contact, would also view reasons other than race as more wrong.

Previous research with minority and majority children found that ethnic minority children were more likely to expect racial exclusion to occur than ethnic majority children (Crystal et al., 2008). In the present study, we theorized that intergroup interaction and cross-race friendships enable children who are in the ethnic majority to become aware of exclusion based on group membership, and this would result in their estimation of how
often racial exclusion occurs. Thus, we expected that European American participants in high-diversity schools would be more aware that race-based exclusion occurs than would European American participants in low-diversity schools, and that this would manifest in their estimations of the frequency of race-based exclusion. A lack of awareness of the prevalence of race-based exclusion may provide an obstacle for viewing such decisions as wrong or unfair; thus, analyses of the estimation of the occurrence of exclusion has implications for the evaluations of how wrong it is to exclude others. Further, it was expected that estimations of exclusion in inter-racial interactions would increase with age, as children become more aware of the existence of prejudice and discrimination (Lerner, 2004).

Method
Participants

Participants were 414 European American children and adolescents in fourth, seventh, and tenth grades, attending 14 public schools, in the suburbs of a mid-size city in the mid-Atlantic region. There were 64 girls and 51 boys in fourth grade ( \( M = 9.82 \) years, \( SD = .43 \)), 102 girls and 70 boys in seventh grade ( \( M = 12.90 \) years, \( SD = .40 \)), and 66 girls and 61 boys in tenth grade ( \( M = 15.90 \) years, \( SD = .38 \)). Students were from schools that were classified as either high (seven schools) or low (seven schools) ethnic diversity. Based on legal and political science (Frankenberg & Orfield, 2007) definitions of school diversity (greater than 20% African American and Latino), we sampled from 14 schools which exceeded the criteria. The ethnic diversity for each school was available as a public document that we obtained prior to data collection. Our high-diversity schools had greater than 25% African American and Latino students enrolled, and our low-diversity schools had less than 15% African American and Latino students enrolled; all high-diversity schools were within the same public school district, and all low-diversity schools were within the same district.

Specifically, the public information for our schools was the following: our two high-diversity elementary schools (with the number of participants that were interviewed from the school in parentheses) had 45% ( \( n = 7 \)) and 42% ( \( n = 27 \)) African American and Latino students enrolled, respectively. The two middle schools had 63% ( \( n = 25 \)) and 29% ( \( n = 47 \)) African American and Latino students enrolled, respectively, and the three high schools had 63% ( \( n = 11 \)), 73% ( \( n = 11 \)), and 48% ( \( n = 12 \)) African American and Latino students enrolled, respectively (for all high-diversity schools, \( M = 52\% \) diversity for enrollment). The three low-diversity elementary schools (with the number of participants from the school in parentheses) had 8% ( \( n = 21 \)), 8% ( \( n = 22 \)), and 2% ( \( n = 37 \)) African American and Latino students enrolled, the three low-diversity middle schools had 13% ( \( n = 29 \)), 10% ( \( n = 43 \)), and 5% ( \( n = 28 \)) African American and Latino students enrolled, and the one low-diversity high school had 10% ( \( n = 76 \)) African American and Latino students enrolled (for all low-diversity schools, \( M = 7\% \) enrolled). All schools from the same district had a shared curriculum.

Our focus was on school diversity composition as the unit of analysis. To test for differences within our “high diversity” category, which included a range of diversity proportions, we analyzed differences for our social reasoning outcomes for the two middle schools that reflected the largest range of diversity (within the “high diversity” category), 63% and 29% ethnic minority enrollment. Using a 2(school: A, B) × 3(stereotype: affirmation, recognition, social contexts of stereotypes) ANOVA, school was not significant, \( F (1, 70) = .017, p > .10 \). Moreover, there were no significant differences for the other schools that had smaller differences regarding the percentage of ethnic representation. Thus, the high-diversity schools were combined together.

There were no district-wide prejudice reduction programs in the school district at the time of data collection. The mean income of the students’ family was middle to low-middle as identified by the school principals, and district reporting of
free-lunch programs, and did not significantly differ across school composition categories.

**Instruments**

**Social reasoning about exclusion** Interviewers administered the social reasoning about exclusion interview, which was audio-taped and later transcribed for coding purposes. Participants were read three stories, each representing a different context in which exclusion typically occurs. The three contexts were: lunch (excluding a friend from a lunch table at school), dance (excluding a friend from a high school dance), and sleepover (excluding a friend from a sleepover party at home). The first two stories described peer-initiated exclusion in the school setting and the third story described peer-initiated exclusion in the home setting.

**Assessments** After each story, participants were asked to explain why exclusion occurred and then respond to six assessments: 1) **Wrongfulness of race-based exclusion** (is it all right or not all right to exclude based on race?); 2) **Wrongfulness of non-race based exclusion** (is it all right or not all right to exclude based on lack of shared interests/school rivalry/parental unfamiliarity—for lunch, dance, and sleepover respectively); 3) **Wrongfulness of group-functioning-based exclusion** (is it all right or not all right to exclude because someone will not “fit in” with the group?); 4) **Stereotype assessment** (responses to an open-ended question about interracial encounters: what is it about race that makes people uncomfortable?); 5) **Estimations of race-based exclusion** (estimates of the frequency of exclusion among peers based on race “how often do you think kids your age might not invite someone to lunch because they are a different race?”); and 6) **Estimations of non-race based exclusion** (“How often do you think kids your age might not invite someone to lunch because they do not share the same interests?”). Responses for the wrongfulness of exclusion rating assessments ranged from 1 (“very, very good”) to 8 (“very, very bad”). Responses for the two-frequency estimation rating assessments ranged from 1 (“never”) to 5 (“always”). Thus, three types of assessments were administered to measure participants: 1) spontaneous **stereotypes** about interracial exclusion, 2) **wrongfulness ratings** about interracial exclusion, and 3) estimations of the frequency of how often interracial exclusion occurs.

**Stereotype assessment coding** Responses to the stereotype assessment were coded with a coding system that included three categories: 1) Stereotype affirmation (affirming or endorsing stereotypes); 2) Stereotype recognition (recognizing that others use stereotypes); and 3) Social contexts of stereotypes (e.g., socialization expectations, a history of segregation that contributes to the manifestation of stereotypes) based on extensive pilot data and previous research on stereotypic responses to exclusion scenarios (Killen et al., 2002). The coding system provided explicit definitions for each category with examples.

Coders read participants’ responses to the stereotype assessment and for each response assigned a number for each of the three categories: 0 = no use of the category, .5 = partial use of the category, and 1.0 = sole use of the category (see Killen et al., 2002; Smetana, 2006; Wainryb, Shaw, Laupa, & Smith, 2001, for an explanation of this methodology). Partial use of the category was used for double-coding, that is, when a participant gave two codable reasons, each code was assigned .5.

Thus, a three-point ordinal scale was used for coding the responses. As an example, if the participant responded to the stereotype assessment by stating that: “Stereotypes have been around a long time, and it’s just wrong to use them, but it’s in the media and it influences people,” then this response would be coded as .5 for stereotype recognition, and .5 for social contexts of stereotypes (and 0 for stereotype affirmation) for this participant. If a participant stated: “Black people like different things so you don’t know what they are going to do” then this would be coded as 1.0 for stereotype affirmation (and 0 for the other two codes). This way, proportions of responses were computed for the stereotype assessment categories and analyses were conducted on these proportions,
based on prior research studies using this data analytic procedure with interview data (Wainryb et al., 2001). We used ANOVA-based statistical tests to analyze these types of data due to our repeated measures design (which are not easily analyzed using other approaches such as log-linear), and the fact that ANOVAs are robust to the problem of empty cells, whereas other data analytic procedures (e.g., log-linear models) necessitate cumbersome data manipulation to address the empty cells issue (see Posada & Wainryb, 2008, for a complete explanation and justification of this data analytic approach).

Reliability coding Coding of the stereotype assessment was conducted by trained graduate research assistants blind to the hypotheses, and conducted on 25% of the transcripts. Cohen's kappa coefficient = .83 (percentage agreement = 89%).

Cross-race friendships in each school environment
Following the social reasoning interview, all students were asked three questions about cross-race friendships from the intergroup contact questionnaire (ICQ) (see Crystal et al., 2008). The ICQ was based on an instrument developed by Orfield (as the leader of a project with other authors) (Kurlaender & Yun, 2001; Orfield, 2001) to verify that intergroup contact in the high-diversity schools was positive. The questions were: 1) At school how many friends do you have who are from a different racial or ethnic group than you?; 2) Outside of school how many friends do you have who are from a different racial or ethnic group than you?; and 3) How many of your friends from your neighborhood are from a different racial/ethnic group than you? Responses to these items ranged from 1 ("none") to 4 ("many").

Dependent measures
Thus, our dependent measures were: 1) Stereotype responses (affirmation, recognition, social contexts of stereotypes); 2) Wrongfulness of exclusion ratings (for race-based exclusion, non-race based exclusion, and group-functioning-based exclusion); and 3) Estimation of frequency of exclusion ratings (for race-based exclusion and for non-race based exclusion). These three dependent measures were analyzed for each of three social contexts that reflected a within-subjects variable as all participants evaluated all three scenarios: lunch (friendship), dance (school dance), and sleepover (birthday party sleepover), for a total of nine assessments. In addition, we had three cross-race/ethnic friendship dependent measures: 1) in-school, 2) out-of-school, and 3) the neighborhood. Our independent variables were age (fourth, seventh, and tenth grades) and school composition (high versus low diversity).

Results
Analyses for the stereotype assessments

Stereotype affirmation A 3 (grade: fourth, seventh, tenth) × 2(school composition: low diversity, high diversity) × 3(context: lunch, sleepover, dance) ANOVA with repeated measures on the last variable was conducted on the proportion of
stereotype affirmation responses. As predicted, analyses revealed a main effect for school composition, $F(1, 395) = 5.60, p = .02, \eta^2 = .01$, grade, $F(2, 395) = 3.22, p = .04, \eta^2 = .02$, and context, $F(2, 790) = 4.11, p = .017, \eta^2 = .02$ (see Table 1 for means). Follow-up $t$-tests with appropriate Bonferroni corrections indicated that, consistent with hypotheses, participants used more stereotypes in the low-diversity schools ($M = .27$) than in the high-diversity schools ($M = .19$). In terms of grade, further analyses showed that, also in line with predictions, the fourth ($M = .26$) and seventh graders ($M = .26$) used more stereotypes than did the tenth graders ($M = .18$). In regard to context, participants used stereotypes to explain racial discomfort in the lunch ($M = .27$) and sleepover ($M = .24$) contexts more than in the dance ($M = .19$) contexts, $p$ values < .01.

### Stereotype recognition

A 3(grade) $\times$ 2(school composition: low diversity, high diversity) $\times$ 3(context: lunch, sleepover, dance) ANOVA with repeated measures was conducted on the proportion of stereotype recognition responses (e.g., the recognition that individuals sometimes use stereotypes to justify exclusion), and revealed a significant main effect for grade, $F(2, 395) = 7.33, p = .001, \eta^2 = .04$, and context, $F(2, 790) = 7.70, p = .001, \eta^2 = .02$ (see Table 1 for means). Across all three contexts, older participants (tenth grade, $M = .33$) recognized stereotypes more than did younger participants (fourth and seventh graders, $Ms = .20, .22$ respectively), $p$ values = .001. Moreover, participants were more likely to recognize stereotypes in the lunch context than in the dance and sleepover contexts ($Ms = .30, .25, .19$ respectively), $p$ values < .001, .02 respectively.

### Social contexts of stereotypes

A 3(grade) $\times$ 2(school composition) $\times$ 3(context: socialization, history of segregation) ANOVA with repeated measures was conducted on the proportion of justifications that referred to the social contexts of stereotypes (e.g., socialization, history of segregation). A main effect of context, $F(2, 790) = 28.88, p = .001, \eta^2 = .06$, and a significant interaction effect of grade by context, $F(4, 790) = 3.68, p = .01, \eta^2 = .02$, emerged.
(see Table 1 for means). Follow-up tests indicated that participants were most likely to refer to the social contexts of stereotypes when evaluating exclusion in the dance context \((M = .56)\) than in the lunch \((M = .33)\) or sleepover \((M = .40)\) contexts. Further analysis of the grade by context interaction revealed significant grade-related differences for references to the social context of stereotypes only in the sleepover context, \(F(2, 407) = 6.83, p = .001, \eta^2 = .03\). Pairwise comparisons showed that the tenth graders \((M = .52)\) referred to the social contexts of stereotypes when evaluating exclusion in the sleepover context significantly more than seventh graders \((M = .41)\) who also referred to this more often than fourth graders \((M = .29)\); there were no significant grade differences for the lunch and dance contexts regarding the proportion of statements referring to the social contexts of stereotypes.

**Wrongfulness of exclusion**

The next set of analyses pertained to evaluations of the wrongfulness of exclusion in interracial contexts as a function of school composition and age (grade) for each of three contexts: lunch, dance, and sleepover (refer to Table 2 for means). Participants were asked to rate the wrongfulness of three different types of exclusion: 1) race-based, 2) non-race-based, and 3) group-functioning-based exclusion. A 3(grade) × 2(school composition) × 3(context) × 3(type of exclusion) ANOVA with repeated measures on the last two factors was conducted, resulting in significant main effects for school composition, \(F(2, 403) = 9.25, p = .003, \eta^2 = .02\) context, \(F(2, 806) = 160.37, p = .001, \eta^2 = .29\), and type of exclusion, \(F(2, 806) = 762.46, p = .001, \eta^2 = .65\).

As hypothesized, participants in the high-diversity schools \((M = 6.30)\) rated all types of exclusion as significantly more wrong than did students in the low-diversity schools \((M = 6.13, p = .003)\). Confirming our hypothesis about parental influence, follow-up analyses revealed that participants rated all types of exclusion as significantly less wrong in the sleepover context \((M = 5.80)\) than either in the friendship context \((M = 6.40, p = .001)\) or in the dance context \((M = 6.40, p = .001)\) or in the dance context \((M = 6.40, p = .001)\). Additionally, confirming hypotheses about the wrongfulness ratings, race-based exclusion was viewed as the most wrong \((M = 7.31)\), followed by group-functioning-based exclusion \((M = 5.95)\), and then non-race-based exclusion, which was rated as the least wrong \((M = 5.42; all p values < .001)\).

In addition, there was a significant interaction between type of exclusion and grade, \(F(4, 806) = 24.00, p = .001, \eta^2 = .11\). Follow-up repeated measures ANOVAs with type of exclusion across the three contexts as the repeated factor confirmed our age-related hypotheses. Specifically, with age, group-functioning-based \((Ms = 6.29, 5.97, 5.50\) for fourth, seventh, and tenth graders respectively) and non-race-based \((Ms = 5.73, 5.49, 4.99\) for fourth, seventh, and tenth graders respectively) reasons for exclusion were viewed as increasingly less wrong (all comparisons, \(p < .05\)). At the same time, however, with age, race-based exclusion was viewed as more wrong by seventh and tenth graders than by fourth graders \((p < .001)\), with no difference between ratings for the seventh and tenth grade participants \((Ms = 7.02, 7.45, 7.39\) for fourth, seventh, and tenth graders, respectively).

Thus, the hypotheses about school diversity were confirmed; participants in high-diversity schools were more likely to view interracial exclusion as more wrong than were participants in low-diversity schools. Not surprisingly, race-based exclusion was viewed as more wrong than exclusion based on group-functioning or non-race-related reasons. Exclusion to make the “group work well” (group functioning) was viewed as more wrong than the other non-race based reasons. Moreover, with age, children viewed race-based exclusion as increasingly more wrong, and non-race based exclusion as increasingly less wrong.

**Estimations of frequency of exclusion**

A 3(grade) × 2(school composition) × 3(context) × 2(reasons for exclusion: non-race based, race-based) ANOVAs with repeated measures on the last two factors were performed on participants’ estimations of the frequency of exclusion.
Table 2. Wrongfulness of exclusion ratings for three contexts by grade and school composition

<table>
<thead>
<tr>
<th>Grade by school composition</th>
<th>N</th>
<th>Context by type of exclusion</th>
<th>Lunch</th>
<th>Race</th>
<th>Non-race</th>
<th>Group functioning</th>
<th>Dance</th>
<th>Race</th>
<th>Non-race</th>
<th>Group functioning</th>
<th>Sleepover</th>
<th>Race</th>
<th>Non-race</th>
<th>Group functioning</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fourth grade</td>
<td>115</td>
<td>Lunch</td>
<td>7.04</td>
<td>5.84</td>
<td>6.51</td>
<td>7.13</td>
<td>6.41</td>
<td>6.28</td>
<td>6.41</td>
<td>6.28</td>
<td>6.89</td>
<td>4.95</td>
<td>6.03</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Dance</td>
<td>9.13</td>
<td>1.05</td>
<td>1.18</td>
<td>7.03</td>
<td>6.29</td>
<td>6.23</td>
<td>7.03</td>
<td>6.23</td>
<td>6.86</td>
<td>5.11</td>
<td>5.91</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sleepover</td>
<td>1.04</td>
<td>1.30</td>
<td>1.23</td>
<td>0.98</td>
<td>1.22</td>
<td>1.07</td>
<td>0.96</td>
<td>1.07</td>
<td>1.04</td>
<td>1.18</td>
<td>1.18</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Low</td>
<td>7.99</td>
<td>6.80</td>
<td>6.51</td>
<td>7.03</td>
<td>6.29</td>
<td>6.23</td>
<td>7.03</td>
<td>6.23</td>
<td>6.86</td>
<td>5.11</td>
<td>5.91</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>High</td>
<td>7.17</td>
<td>5.94</td>
<td>6.51</td>
<td>7.37</td>
<td>6.69</td>
<td>6.40</td>
<td>7.37</td>
<td>6.40</td>
<td>6.94</td>
<td>4.57</td>
<td>6.29</td>
<td></td>
</tr>
<tr>
<td>Seventh grade</td>
<td>171</td>
<td>Lunch</td>
<td>7.51</td>
<td>5.77</td>
<td>6.16</td>
<td>7.51</td>
<td>6.05</td>
<td>5.89</td>
<td>7.51</td>
<td>6.05</td>
<td>7.32</td>
<td>4.64</td>
<td>5.84</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Dance</td>
<td>7.50</td>
<td>5.59</td>
<td>6.05</td>
<td>7.51</td>
<td>5.95</td>
<td>5.74</td>
<td>7.51</td>
<td>5.95</td>
<td>7.35</td>
<td>4.46</td>
<td>5.59</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sleepover</td>
<td>7.54</td>
<td>6.01</td>
<td>6.32</td>
<td>7.51</td>
<td>6.18</td>
<td>6.11</td>
<td>7.51</td>
<td>6.18</td>
<td>7.28</td>
<td>4.90</td>
<td>6.18</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Low</td>
<td>100</td>
<td>5.69</td>
<td>1.08</td>
<td>6.63</td>
<td>0.99</td>
<td>1.12</td>
<td>6.63</td>
<td>0.99</td>
<td>0.76</td>
<td>1.36</td>
<td>1.34</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>High</td>
<td>1.09</td>
<td>5.71</td>
<td>1.02</td>
<td>1.04</td>
<td>1.10</td>
<td>1.09</td>
<td>1.04</td>
<td>1.10</td>
<td>1.06</td>
<td>1.48</td>
<td>1.32</td>
<td></td>
</tr>
<tr>
<td>Tenth grade</td>
<td>127</td>
<td>Lunch</td>
<td>7.50</td>
<td>5.20</td>
<td>5.76</td>
<td>7.39</td>
<td>5.52</td>
<td>5.37</td>
<td>7.39</td>
<td>5.52</td>
<td>7.26</td>
<td>4.24</td>
<td>5.38</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Dance</td>
<td>7.38</td>
<td>5.04</td>
<td>5.71</td>
<td>7.28</td>
<td>5.51</td>
<td>5.16</td>
<td>7.28</td>
<td>5.51</td>
<td>7.24</td>
<td>4.22</td>
<td>5.29</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sleepover</td>
<td>7.67</td>
<td>5.45</td>
<td>5.82</td>
<td>7.30</td>
<td>5.54</td>
<td>5.68</td>
<td>7.30</td>
<td>5.54</td>
<td>7.30</td>
<td>4.26</td>
<td>5.52</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Low</td>
<td>.71</td>
<td>1.14</td>
<td>1.14</td>
<td>.67</td>
<td>1.12</td>
<td>1.24</td>
<td>.67</td>
<td>1.12</td>
<td>.85</td>
<td>1.32</td>
<td>1.30</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>High</td>
<td>.75</td>
<td>1.08</td>
<td>1.10</td>
<td>.78</td>
<td>0.945</td>
<td>1.23</td>
<td>.78</td>
<td>0.945</td>
<td>.78</td>
<td>1.17</td>
<td>1.21</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>.62</td>
<td>1.20</td>
<td>1.20</td>
<td>.95</td>
<td>1.36</td>
<td>1.20</td>
<td>.95</td>
<td>1.36</td>
<td>.95</td>
<td>1.54</td>
<td>1.43</td>
<td></td>
</tr>
</tbody>
</table>

Note: N = 413. Non-race = Non race-based exclusion (for friendship = peer lacks shared interests; for dance = peer attends rival school; for sleepover = parental unfamiliarity with peer). Race = race-based exclusion (due to race of peer). Group functioning = “lack of fit” with the group.
Supporting our hypotheses, a main effect was found for grade, $F(2, 406) = 3.39, p = .035$, $\eta^2 = .016$, indicating that fourth graders ($M = 2.56$) estimated less exclusion overall than did the seventh graders ($M = 2.65$) or tenth graders ($M = 2.73$), $p = .01$ (see Table 3 for means). An interaction between context and grade was found as well, $F(2, 812) = 10.13, p = .001$, $\eta^2 = .048$. Follow-up univariate ANOVAs indicated the interaction was driven by a grade effect in the sleepover context, with tenth graders estimating more overall exclusion than either fourth or seventh graders ($p$ values < .01). There were no grade effects for the dance context, $F(2, 409) = 2.60, p = .076$, $\eta^2 = .013$.

Confirming our hypotheses, main effects were found for type of exclusion, $F(2, 406) = 477.65, p = .001$, $\eta^2 = .54$, indicating that participants estimated non-race-based exclusion ($M = 3.04$) to be more frequent than race-based exclusion ($M = 2.25$), $p = .001$. A main effect for the context of exclusion, $F(2, 812) = 17.66, p = .001$, $\eta^2 = .042$, revealed that estimations of exclusions for the dance context were less frequent ($M = 2.52$) than either for lunch ($M = 2.73$), or for sleepover ($M = 2.69$), $p = .001$. Additionally, there was an interaction between context and type of exclusion, $F(2, 812) = 63.440, p = .001$, $\eta^2 = .14$. Follow-up paired $t$-tests indicated that participants estimated that non-race-based exclusion occurs more frequently than race-based exclusion in the lunch ($M = 3.23$ vs. 2.25) and sleepover ($M = 3.22$ vs. 2.12) contexts ($p$ values < .001). Participants did not differentiate between the two types of exclusion in the dance context ($Ms = 2.69$ vs. 2.36).

### Cross-race friendships

A 3(grade: fourth, seventh, tenth) $\times$ 2(school composition: high diversity, low diversity) $\times$ 3(context
of friendship measures: in-school, out-of-school, neighborhood) repeated measures ANOVA with repeated measures on the last factor revealed a main effect for school composition, \( F(1, 403) = 59.00, \ p = .001, \ \eta^2 = .13 \), such that enrollment in high-diversity schools (\( M = 3.00 \)) reflected a higher level of self-reported cross-race friendships than enrollment in low-diversity schools (\( M = 2.50 \)). A main effect also emerged for the context of friendships, \( F(2, 806) = 16.3, \ p = .001, \ \eta^2 = .29 \). Participants reported more cross-race/ethnic friends in school (\( Ms = 3.30 \)) than out-of-school (\( Ms = 2.90 \)) and fewer in their neighborhood (\( Ms = 2.30 \)).

Discussion

In this study, novel findings pertained to the role of ethnic diversity of the school, the age of the participant, and the context of exclusion. Each of these findings will be discussed in turn.

School diversity

Our study, conducted in the suburbs of Washington, DC, supported the view that ethnic diversity has the potential to have a positive benefit for all students, including European American children and adolescents (Frankenberg & Orfield, 2007; Killen, Crystal, & Ruck, 2007). We found that European American students attending ethnically-diverse schools were more likely to have cross-group friendships than European American students attending schools with little or no ethnic diversity. While it was not surprising that children at schools with low diversity did not have cross-group friendships (given the absence of opportunity), it is important to document the extent to which participants attending ethnically-diverse schools did have cross-group friendships. This is because school diversity (contact) alone does not result in intergroup friendships. For example, tracking and other school infrastructures, such as magnet schools, often create ethnic segregation within ethnically-diverse schools (Frankenberg & Orfield, 2007). In this study, however, students attending ethnically-diverse schools had more cross-group friendships than did students attending schools with little ethnic diversity.

The main finding was that participants enrolled in heterogeneous schools used fewer stereotypes to explain interracial exclusion. Moreover, they also were more likely to evaluate interracial exclusion as wrong. These findings support intergroup contact theory which has been studied extensively in adulthood (Hewstone et al., 2002) but rarely in childhood (for an exception, see Tropp & Prenovost, 2008). In research with adults the positive effects of intergroup contact have been measured using general scales such as favorability of an outgroup member. In contrast, this research focused on how experiences in school are related to the evaluation of exclusion in familiar everyday peer relationships. Second, these findings add to a growing body of evidence demonstrating why diversity may be beneficial for all students. These findings support school desegregation policy initiatives, which have typically used external indices such as increased grade point averages and lower delinquency rates to measure positive outcomes of desegregation (Frankenberg & Orfield, 2007; Kurlaender & Yun, 2001). What makes these findings novel is that the outcomes reported in this study reflect psychological data, particularly regarding social cognition and evaluations of interracial interactions. Moreover, few studies on the value of diversity have examined the ways in which diversity may provide a positive developmental benefit for majority European American students (for an exception, see Tropp & Prenovost, 2008). Recently, Page-Gould et al. (2008) found that adult college students who revealed implicit racial biases, and experienced stress in interracial contexts, were less likely to experience anxiety in these encounters after they made a cross-race friendship. Using a diary method, Page-Gould and her colleagues were able to demonstrate a direct link between having a cross-race friendship and less anxiety as measured by cortisol samples. This was found for both ethnic majority and minority students at the college-age level. Further, Mendes, Gray, Mendoza-Denton, Major, and Epel (2007) reported negative physiological effects related to stressful intergroup encounters, indicating that...
overcoming this stress is important for positive interactions in school and the workforce. Understanding the developmental processes through which cross-race friendships help to lessen anxiety in interracial encounters would be an important extension of the current study.

Other factors also influence the relationship between intergroup contact and racial attitudes, and should be investigated. For example, children in heterogeneous schools with intergroup contact may also be exposed to more positive messages about intergroup relationships, and may also attend these schools as a result of parental decisions to expose their children to diverse environments. Thus, parental attitudes and school environment, among other variables, remain to be investigated to fully understand why intergroup contact contributes to more positive racial attitudes.

**Age-related findings**

With age, children judged it increasingly wrong to exclude on the basis of race in interracial peer contexts. This finding is consistent with prior studies on straightforward racial exclusion (Killen et al., 2002) as well as research on adolescents' concepts of discrimination and rights (Helwig, 2006). At the same time, we also found that, with age, participants judged interracial exclusion to be increasingly less wrong when the reason given for exclusion by the excluder reflected non-race-based reasons (e.g., a lack of shared interests, parental unfamiliarity, or membership in a rival high school). This may be because non-race-based reasons are viewed as legitimate in the context of friendship given that friendship is typically regarded by adolescents as a matter of personal choice (Nucci & Turiel, 2000). From this perspective, reasons such as lack of shared interests or unfamiliarity could be perceived as a valid basis for rejecting a friend or excluding a peer. Killen, Henning, et al. (2007) found that ethnic minority participants were more likely than ethnic majority students to be critical of non-race-based reasons given by European American students for excluding a minority student from a friendship encounter (with no reported age differences). The authors speculated that minority students viewed the purportedly non-race-based justifications, in fact, as proxies for race-based exclusion. These different attributions about cross-race encounters have the potential to contribute to the decline of cross-race friendships with age (Aboud, Mendolsohn, & Purdy, 2003), and thus are cause for concern and more in-depth investigation.

What also remains to be researched more thoroughly are the age-related reasons that individuals provide for justifying exclusion in interracial interactions. What types of reasons serve as a legitimate basis for excluding someone of a different race? In this study, due to time constraints with the administration of the instruments, only one non-race-based reason was described to participants per context (e.g., lack of shared sports interests in the lunch context). Investigating how individuals evaluate a range of reasons would shed much light on this aspect of the decision-making process.

Moreover, the use of stereotypes to explain racial discomfort decreased with age, while the recognition that others use stereotypes to justify exclusion increased with age. The decrease in the use of stereotypes is consistent with other literature showing that young children are more likely to explicitly articulate racial stereotypes than are older children (Nesdale, 2004). Yet stereotypes were not absent in older children given that school diversity and opportunities for contact were significantly related to the use of stereotypes in the adolescent samples. Further, with age, participants were more likely to estimate that racial exclusion occurs among their peers. Thus, with age, participants were aware that stereotypes are used to justify interracial peer exclusion, and that this type of exclusion occurs with some frequency.

These age-related findings are novel and provide new evidence about the developmental trajectory of racial attitudes. Prior studies revealed that stereotypes often diminish soon after early childhood, around seven or eight years of age. In this study, though, adolescents, with low contact, used stereotypes to explain discomfort with racial interactions. The measure employed in this study differed from prior studies with younger children,
which measured the attribution of traits (nice, mean) to individuals based on skin color (Aboud & Amato, 2001). Asking adolescents why it is that individuals experience racial discomfort is a more subtle and nuanced method for assessing the use of stereotypes. Research has shown that adults implicitly associate negative traits to individuals based on skin color (Banaji, Baron, Olson, & Dunham, 2008), and the method used in the present study provides a way of examining explicit stereotypes but with a more indirect measurement than the one used in prior studies with children.

**Context**

Our findings revealed different patterns of responses across the three contexts, supporting social psychology findings on differences for racial bias as a function of the social context (Wittenbrink, Judd, & Park, 2001). Wittenbrink et al. (2001) demonstrated that racial attitudes vary when participants are shown minority individuals in negative (graffiti scene) or positive (family picnic) contexts. Further, Wainryb (2006) has argued for contextual approaches when investigating social and moral evaluations in childhood. In this study, participants were asked to evaluate three contexts of interracial exclusion that reflected three types of intimate relationships (having lunch with someone, inviting them to a sleepover party, and inviting someone to a school dance), in contrast to work- or academic-related exclusion contexts. This was by design, and was done because it has been documented that intimate relationships remain one of the last contexts to reflect ethnic diversity (Kennedy, 2003). While most individuals reject using race as a reason to not vote for someone for president or to refrain from hiring someone, interracial marriage and dating rates remain quite low in the US (Kennedy, 2003). In this study, the types of exclusion contexts evaluated by participants were ones commonly experienced by ethnic minority students (as indicated in pilot work).

We expected that the home context (sleepover) would be viewed quite differently from the other school-based contexts. Regarding stereotype usage, this expectation was not borne out. In fact, participants used stereotypes to explain racial discomfort for the lunch and sleepover context to the same degree (no significant difference) and used stereotypes for these contexts more than for explaining racial discomfort at the school dance. When evaluating the school dance, participants frequently referred to the context of stereotypes in terms of the social and historical factors that might account for discomfort and were much less likely to evoke explicit stereotypes, as was revealed in the lunch and sleepover contexts. Thus, participants viewed the dance context as more a matter of larger societal expectations that contribute to interracial exclusion than as individual stereotyping that occurs in situations such as not having lunch with someone or not inviting them over to their house for a sleepover party.

When asked how wrong it was to exclude, however, participants were more likely to condone exclusion at home than in the school contexts. Thus, the results for the wrongfulness of exclusion ratings (but not the stereotypes usage) supported our expectation that exclusion would be more condoned in the home context than in the other two contexts.

**Conclusion**

These findings show a complex pattern of judgments, and indicate that children and adolescents make context-sensitive decisions about interracial exclusion. The findings also suggest that interracial exclusion in familiar peer contexts in childhood remains a societal problem to address. Students attending schools in a metropolitan region of the US continue to espouse stereotypes about race through adolescence (between 20% and 30% in the present sample). The positive aspects of the findings in this study were that attending diverse schools with positive intergroup contact (as measured by cross-group friendships) helped to lessen the use of negative racial attitudes, as held by ethnic majority children and adolescents. Reducing intergroup bias and prejudice is a necessary goal for a just and fair society (Cameron, Rutland, & Brown, 2007; Killen & Smetana, in press).
Understanding how biases emerge in development is essential for creating effective programs to reduce bias, and to inhibit the formation of stereotypic expectations, which have been directly linked to prejudicial and discriminatory behavior in a wide range of social contexts (Aboud & Levy, 2000; Abrams et al., 2005; Dovidio & Gaertner, 2004; Gaertner & Dovidio, 1986; Killen, 2007; Levy & Killen, 2008; McKown, 2005).

Acknowledgments

The research described in this paper was supported by a grant from the National Institute of Child Health and Human Development (1R01HD04121-01). We thank the research team at the University of Maryland, Marguerite Adams, Holly Bozeman, Alaina Brenick, Michael J. Collins, Christina Edmonds, Julia Hadrichy, Alexandra Henning, Nancy Geyelin Margie, Heidi McGlothlin, Alexander O’Connor, Christine Pitocchelli, and Stefanie Sinno; and the research team at Georgetown University, Nancy Gibbs, Lindsey Gansen, Dayna McGill, Elizabeth Kravec, Anne Blossom, Mia Shorteno-Fraser, and Meredith Mellody for assistance with instrument development, data collection, data coding, analyses, and reliability coding. In addition, we extend a special thanks to Holly Bozeman for assistance with the stereotype measure and the stereotype coding system. Appreciation is extended to the children who participated in this study, and to the parents who gave consent as well as to the staff and school districts for their participation.

References


Edmonds, C., & Killen, M. (2009). Do adolescents’ perceptions of parental racial attitudes relate to their intergroup contact and cross-race relationships? Group Processes & Intergroup Relations, 12, 5–21.


**Biographical notes**

**MELANIE KILLEN** is Professor of Human Development, Affiliate Professor of Psychology, and Associate Director of the Center for Children, Relationships, and Culture at the University of Maryland. Her research interests include social and moral reasoning, cultural influences on development, social reasoning about group inclusion and exclusion, moral evaluations of intergroup bias, prejudice in childhood, intergroup bias, and how intergroup attitudes are related to judgments about group identity and fairness.
MEGAN CLARK KELLY is a doctoral candidate and lecturer in the Department of Human Development at the University of Maryland. Her interests are exclusion, gender role stereotypes, and morality in childhood.

CAMERON RICHARDSON is a doctoral candidate and lecturer in the Department of Human Development at the University of Maryland. His interests are exclusion, morality, theory of mind, social cognitive development, and stereotyping.

DAVID CRYSTAL is Associate Professor at the Department of Psychology at Georgetown University. He has wide-ranging interests in developmental, clinical, and cross-cultural psychology. He is especially concerned with how cultural contexts affect such psychological phenomena as tolerance of difference, concepts of normality and deviance, attitudes toward self-change, and disturbances in socioemotional adjustment among children and youth.

MARTIN RUCK is Associate Professor of Urban Education and Psychology at the Graduate Center at the City University of New York. His research examines the overall process of cognitive socialization—at the intersection of race, ethnicity and class—in terms of children’s thinking about human rights and social justice.