

Gender Similarities and Differences in Children's Social Behavior: Finding Personality in Contextualized Patterns of Adaptation

Audrey L. Zakriski
Connecticut College

Jack C. Wright
Brown University

Marion K. Underwood
University of Texas at Dallas

This research examined how a contextualist approach to personality can reveal social interactional patterns that are obscured by gender comparisons of overall behavior rates. For some behaviors (verbal aggression), girls and boys differed both in their responses to social events and in how often they encountered them, yet they did not differ in overall behavior rates. For other behaviors (prosocial), gender differences in overall rates were observed, yet girls and boys differed more in their social environments than in their responses to events. The results question the assumption that meaningful personality differences must be manifested in overall act trends and illustrate how gender differences in personality can be conceptualized as patterns of social adaptation that are complex and context specific.

Keywords: gender differences, personality assessment, aggression, withdrawal

Over the last few decades, explanations of gender differences in social behavior have increasingly focused on context. Social interactive models have highlighted the variability of gender-related behavior and traced gender differences (and similarities) to interpersonal dynamics and situational demands (Deaux & Major, 1987). Recently, the *two cultures* view has suggested that girls and boys differ in their social behavior largely because their sex-segregated peer groups elicit behaviors that may not be characteristic of them in other social contexts (Maccoby, 1998, 2002). These and related views have clarified why gender differences are variable over studies and have identified social forces that might contribute to differences previously assumed to be inherent or “essential” (Archer, 1996; Eagly, 1995; Leaper, 2000).

In place of a few stable, enduring, and broad gender differences in behavior, what has emerged is a mosaic consisting of small differences in some contexts, no differences in others, and “reversed” differences in others. As the mosaic’s complexity grows, the usefulness of studying gender differences in personality seems to recede (see Leaper, 2000; Maccoby, 1998). Indeed, some have

debated whether we should stop studying such differences altogether (Baumeister, 1988; Eagly, 1987; Lott & Eagly, 1996). We suggest that the study of personality and context need not conflict; rather than receding as the mosaic grows, personality is revealed *in* the mosaic itself. In our view, legitimate criticisms of research on gender differences in personality arise more from the limitations of acontextual trait models than from the concept of personality per se.

Studies of gender and personality have often focused on overall behavior rates, average trait ratings, or summary checklist scores (see meta-analyses by Feingold, 1994). This is consistent with other research on individual differences in which behavior inventories attempt to “filter out” situational variation to assess the “stable and enduring” features of the person (Barkley, 1988). One of the clearest theoretical justifications for such methods can be found in the *act frequency* approach to personality (Buss & Craik, 1983), which defines a disposition as an *act trend*, or the number of relevant acts a person displays over a period of observation. Although contemporary theorizing about personality often embraces more complex and dynamic views of traits (Johnson, 1999; Roberts & Caspi, 2001), the predominance of act frequency methods perpetuates the notion that “consistency across situations lies at the core of the concept of personality” (Weiten, 2004, p. 478; see also Gray, 2002). It is the clash between this concept of personality and the context dependence of gender differences in aggression, affiliation, helping, and other behaviors (Archer, 2000; Caldera & Sciaraffa, 1998; Eagly & Crowley, 1986) that has led some to reject the notion of gender differences in “underlying traits or abilities” (Leaper, 2000, p. 391).

There have long been alternative views of personality that more explicitly incorporate context. G. W. Allport (1937) discussed the need to understand how contexts elicit behavior and advocated idiographic approaches that could reveal response styles that distinguish one individual from another. F. Allport (1924) argued that

Audrey L. Zakriski, Department of Psychology, Connecticut College; Jack C. Wright, Department of Psychology, Brown University; Marion K. Underwood, School of Behavior and Brain Sciences, University of Texas at Dallas.

We thank the children, staff, and families whose cooperation made it possible to collect the data reported here. We are deeply grateful to Harry Parad, Director of Wediko Children’s Services, whose support made this project possible, and the numerous research coordinators and assistants involved over the years. Special acknowledgment is due to the late Hugh Leichtman, previously Executive Director of Wediko, who facilitated research at Wediko over many years.

Correspondence concerning this article should be addressed to Audrey L. Zakriski, Department of Psychology, Connecticut College, 270 Mohegan Avenue, New London, CT 06320. E-mail: alzak@conncoll.edu

personality is best understood as the quality of a person's "adaptation to the social features of the environment" (p. 149) and spoke of the process by which a person's responses can be reevoked or increased by the reactions those responses elicit from others. In a similar spirit, a contextualist approach to personality suggests that the quality of a person's adaptation can be seen in the patterning of *if . . . then* links between contexts and behaviors (Wright & Mischel, 1987; see also Mischel & Shoda, 1995; Vansteelandt & Van Mechelen, 1998). Rather than reflecting noise to be averaged out, contextual variability is viewed as an inherent feature of personality, one that can be used to clarify what is distinctive about individuals or groups and to probe how people describe and interpret behavior (e.g., Dawson, Zeitz, & Wright, 1989; Shoda, Mischel, & Wright, 1989).

Research on anxiety (Van Mechelen & Kiers, 1999), hostility (Vansteelandt, 1999), neuroticism, and other traits (Bolger & Schilling, 1991; Van Heck, Perugini, Caprara, & Fröger, 1994) has shown how the study of reactions in context can clarify the situational specificity of individual differences in these domains. Other work has illustrated how contextual approaches can enhance clinical diagnosis (Gresham & Noell, 1993; Scotti, Morris, McNeil, & Hawkins, 1996). In child assessment, groups defined as primarily aggressive, primarily withdrawn, or both aggressive and withdrawn using a popular acontextual measure have been found to include functional subgroups that differ in the events that elicit their behaviors and in their rates of encountering those events (Wright & Zakriski, 2001). Related research (Wright, Lindgren, & Zakriski, 2001) found that people's personality judgments were sensitive to the contextual origins of children's behavior (e.g., aggression resulting from encountering aversive events or from aggressive reactions to those events), whereas standardized acontextual measures were not.

Conceptualizing personality in terms of contextualized behavior patterns or "signatures" has implications for the study of individuals (Shoda, Mischel, & Wright, 1993); trait groups (Wright & Zakriski, 2001); and, we suggest, gender differences. First, this approach highlights the possibility that girls and boys show different response patterns even when they have similar overall behavior rates. Suppose that Ann is aggressive when warned by adults but not when provoked by peers; Brian is aggressive when provoked but not when warned. If they encountered similar rates of these contexts, these children would have similar overall rates of aggression and would be equally aggressive in an act frequency sense. In a contextualist view, however, the children's distinctive patterns reveal what is so different about their aggressiveness and, more broadly, their personalities.

Second, the contextual approach highlights the possibility that girls and boys show similar response patterns even when they differ in overall behavior rates. Suppose that Sue shows the same pattern of reactions as Brian (aggressive when provoked) but that she is rarely provoked by peers, whereas Brian is often provoked. From an act frequency perspective, Brian is more "aggressive" because of his higher overall rate of such acts. Inferences about the children's personalities, however, could be misguided if their behavior rates stem from social experiences rather than reaction tendencies. Interpreting overall behavior rates becomes even more challenging when both social environments and reaction patterns differ, as may be the case for girls and boys.

Distinctions between the present and related views should be noted. Gender researchers at times use context differences to "explain away" gender differences in personality. For example, on the basis of high provocation rates among boys, it is argued that girls and boys would not differ in aggressiveness if this context effect were removed (see Maccoby, 1998). In our view, this type of analysis is useful but potentially incomplete. Higher provocation rates may contribute to boys' overall frequency of aggression, but boys and girls may still differ in their responses to provocation. Conversely, even if girls and boys were similar in their overall frequencies for a behavior (e.g., withdrawal), they could still differ in how often they encounter eliciting events and in how they react to them. A contextual analysis of behavior is a two-edged sword: It may explain away some gender differences in overall rates of behavior yet also reveal gender differences in response patterns that exist even when children's overall behavior rates are equal.¹

Contemporary theories of personality also consider contextual influences (Caspi & Roberts, 2001; Helson, Jones, & Kwan, 2002; Muller, Endler, & Parker, 1990). Although context is recognized as interacting with traits as personality is shaped over time, too often it is not fully integrated into the conceptualization and measurement of traits themselves. The result is a disparity between theorizing about personality that is richly contextualized and trait measures that in themselves are not (e.g., the Big Five, the California Psychological Inventory, Cattell's 16PF). This reinforces the view that assessment is an atheoretical, actuarial task and that personality is "what personality instruments measure" (Feingold, 1994, p. 429). Our position is closer to Coombs's (1964), who argued that data never "speak for themselves" and that all measurement rests on assumptions about why data should be collected one way as opposed to another. Even at the miniature level of measuring how often a behavior occurs, it is essential to ask what environmental and psychological processes may have contributed to the measurement before drawing conclusions about the "traits" of the individual. Understanding gender differences in personality may require not only that we contextualize our interpretations of aggregated measures after they have been collected but also that we explicitly contextualize the assessment process from the outset.

This study examined gender differences in social behavior using extensive observations of girls and boys at a summer program for children with behavior problems. We examined children's overall rates of behavior (physical and verbal aggression, withdrawal, prosocial behavior), how often they encountered contexts that might elicit these behaviors (e.g., peer talk, adult punish), and how they responded when each context occurred. We tested three propositions. First, we expected differences in overall behavior rates to be consistent with research that used comparable measures. Although specific findings vary, girls generally display less physical aggression and more prosocial behavior than boys (Coie & Dodge, 1998; Eisenberg & Fabes, 1998; Hyde, 1984). Evidence on verbal aggression is mixed: Many studies have found no difference, whereas others have found that girls are more verbally

¹ In the present study, we place no priority on finding gender differences (*alpha bias*; Hare-Mustin & Marecek, 1988) or similarities (*beta bias*). Instead, we suggest that neither can be adequately understood through comparisons of overall frequency measures that ignore the rates of encountering contexts and the patterning of people's responses to them.

aggressive (Archer, Pearson, & Westeman, 1988; Österman et al., 1998). Gender differences have seldom been found for withdrawal (Mullen, Snidman, & Kagan, 1993; Rubin, Burgess, & Coplan, 2002). The effect of age is unclear because of a confounding of age and methodology: Self- and other-reports often used with older children are more influenced by sex role stereotyping than are direct observations often used with younger children (Eisenberg & Fabes, 1998; Rubin et al., 2002). The present study used observational methods across a wide age range in an effort to address this issue.

Second, we tested the proposition that comparisons of overall behavior rates do not directly show differences (or similarities) in the personalities of girls and boys. Rather, overall rates can be understood only in light of the processes that contribute to them, including the quality of children's social environments. Specifically, past research has suggested that gender differences in behavior occur partly because girls are less likely than boys to experience peer provocation, adult warnings, and adult punishments, whereas girls are more likely to experience adult praise and peer talk (Buhrmester & Prager, 1995; Leaper, Anderson, & Sanders, 1998; Smith & Boulton, 1990; Zarbatany, McDougall, & Hymel, 2000). Although we expected social environment differences across the age range we studied (7–15 years), available evidence suggests that girls' and boys' social worlds become more different with age (Maccoby, 1998).

Third, we tested the claim that gender comparisons of overall behavior rates obscure narrow but important differences between girls and boys in the patterning of their reactions to contexts. Research has seldom compared girls' and boys' reactions to multiple contexts within the same study; instead, the contextual variability of gender differences is usually extracted from reviews of multiple studies (Eisenberg & Fabes, 1998; Underwood, 2003). Thus, our predictions extrapolate from indirect evidence. We expected boys (compared with girls) to be especially likely to show physical aggression in response to peer provocation (Archer et al., 1988; Coie & Dodge, 1998). Conversely, we expected girls to show more conflict-reducing (prosocial and withdrawn) behavior to peer provocation (Fabes & Eisenberg, 1992; Miller, Danaher, & Forbes, 1986). Some indirect evidence has suggested that girls are more prosocial than boys in interactions with adults but comparably prosocial in interactions with peers (Brody, 1999; Eisenberg & Fabes, 1998; Kochanska & Aksan, 1995). Other studies have suggested that conflict with adult caretakers may be especially likely to elicit aggression from girls (Hoffenaar & Hoeksma, 2002; Walsh, Pepler, & Levene, 2002).

Method

Our data were from a multiyear project at Wediko Children's Services' 45-day residential summer program for children with behavior problems (i.e., aggression, withdrawal, poor social skills). Wediko admits approximately 150 children each summer, referred primarily from public schools in the Boston area and elsewhere in New England. Children live in groups of 8 to 10 same-sex, same-aged peers and participate in a schedule that includes a range of daily activities (e.g., academics, art, swimming). The girl:boy ratio is approximately 1:3, reflecting higher rates of overt behavior problems and referrals for boys, especially prepubertally (Goodman et al., 1997). Approximately 100 counselors and teachers participated each summer in data collection. Permission to use data for research purposes was obtained from parents/guardians during the interview process.

Child and Adult Participants

Children. Data were obtained on 690 children over five summers. Of these, 624 (90.4% of 690) were in residence for the entire summer and had sufficient observations (see below). All girls with sufficient observations were included ($n = 180$). Including all boys with sufficient observations would have involved tradeoffs. As Hays (1973) noted, increasing sample size is desirable when the goal is to obtain precise estimates of parameters, but caution is needed when the goal is to identify effects that are robust enough to warrant further investigation. Using all boys also would have created a disparity between boys and girls in the power of tests of age differences within gender groups. Therefore, 180 boys were randomly sampled, and the present study examined data for 360 children.² The composition of the sample was 53% White, 37% African American, 7% Hispanic, 1% Asian, and 2% other; the children were predominantly lower and middle socioeconomic status. Children were divided into two age groups using a median cutoff of 11 years: younger girls ($n = 93$; mean age 9 years 2 months), younger boys ($n = 93$; mean age 9 years 4 months), older girls ($n = 87$; mean age 12 years 11 months), older boys ($n = 87$; mean age 13 years 0 months). Wechsler Intelligence Scale for Children—Revised (Wechsler, 1974) scores obtained by Hayes (1995) on a subsample of children indicated that their scores fell in the average range: Verbal ($M = 97.9$), Performance ($M = 99.1$), and Full Scale ($M = 98.5$).

Adults. A total of 519 counselors, teachers, and supervisors were employed at Wediko over the 5-year period; these staff were students or recent graduates recruited from numerous colleges and universities. Teachers and activity counselors ran classes and activities for 4 hr each day; cabin counselors remained with their group most of the day. Four to 6 counselors and a supervisor were assigned to each group of children. More female than male cabin counselors were assigned to girls' groups; male:female staff ratios were roughly equal in boys' groups. Of the 519 staff, 410 nonsupervisory staff provided observational data as described below.

Behavior Coding Materials and Procedure

As in related research (Patterson, 1982), the coding system emphasized behaviors that were observable and common in everyday interactions. Earlier methods used at Wediko (Shoda et al., 1993) were expanded to create a set of 7 *context* codes and 10 *response* codes (Wright et al., 1999). The response codes included aggressive, withdrawn, and prosocial behaviors often sampled in past studies: "hit or physically attacked"; "teased, provoked or ridiculed"; "bossed, bullied, or threatened"; "argued or disapproved"; "withdrew, isolated self"; "whined or cried"; "talked in age-appropriate way"; "attended or listened to other(s)"; "showed positive emotion"; and one item not used here, "self-stimulation/self-abuse." The contexts included positive and aversive adult and child behaviors found in previous research to elicit individual differences for these behaviors: "adult praised the child verbally"; "adult gave the child a warning"; "adult instructed the child to do something"; "adult gave the child a time out"; "peer talked in age-appropriate way"; "peer teased, provoked, or ridiculed"; "peer bossed, bullied, or threatened."

The coding manual provided definitions of each code and explained how to record contexts and responses. Staff training included group presentations, practice coding of written and role-play vignettes, and individual and

² Portions of the data from one of these years were analyzed elsewhere (Wright, Zakriski, & Drinkwater, 1999; Wright & Drinkwater, 1997), but the focus of those studies was on syndromal groups and on undergraduates' judgments of personality. A small number of girls was included in Wright et al. (1999), but behavioral observations in that study were normed within gender group, thereby limiting the exploration of gender differences. The current article uses a multiyear sample not previously reported, focuses on gender and age rather than syndrome groups, and examines prosocial behavior as well as aggression and withdrawal.

group feedback from research staff. Coders observed children during hourly activity and class periods and recorded their behavior at the end of the period. Recordings were made on optical scan forms to minimize memory and writing demands. Each coder typically completed forms for 2 to 6 children per period, and each child was observed in three to five periods each day. Each child was observed by a range of coders, including cabin staff who interacted with the child for at least 8 hr per day, teachers who interacted with the child each day for 1.5 hr, and activity staff who interacted with the child during their activity specialty (e.g., crafts, archery).

At the end of each period, the coder identified the target child and provided a global rating of how often she or he displayed each of the 10 responses ("On the whole, how did the child behave during this observation period?") using a 0–3 scale (*not at all, somewhat, moderately, a lot*). The coder then recorded whether the target encountered each of the seven contexts, the name of the peer or adult involved, and the responses shown to each recorded context. If two or more instances of the same context occurred, coders recorded the most recent one to minimize memory demands. Multiple responses were allowed for a given context.

As in previous work (Wright et al., 1999), categories were formed from individual codes. Nine response codes were combined into four composites: verbal aggression (argue, tease, and boss), physical aggression (hit/physically attack), withdrawal (withdraw and whine), and prosocial behavior (talk, attend, positive emotion). We did not analyze self-stimulate because its frequency was low and because it is not commonly studied in the gender literature. Peer boss/threat and peer tease were combined because of the relatively low frequency of peer threat, yielding six context categories: adult praise, adult instruct, adult warn, adult punish, peer talk, and peer boss/tease.

"Overall behavior rates" were computed for each response category by averaging the mean behavior ratings a child received over the summer. "Context rates" were computed by dividing the number of times each child encountered each context by the total number of contexts encountered. "Reaction rates" were computed by dividing the number of responses to a context by the number of times that context occurred (e.g., the number of instances of withdraw in response to boss/tease divided by the number of instances of boss/tease). This yielded a matrix of 4 (response) \times 6 (context), or 24 reactions. For our main analyses, each measure was converted to a z score—that is, $(\text{raw score} - M_{\text{all children}})/SD$ —yielding variables with $M = 0$ and $SD = 1$. A positive (negative) z indicates that a group was high (low) relative to the sample as a whole; a z near 0 indicates the group was near the mean for the sample.

The sample of 360 children had a mean per child of 115.8 hourly observations, 386.6 responses to contexts, and 215.0 instances of contexts. Context frequencies (with mean relative frequencies) per child per context were as follows: adult instruct 64.7 (.30), adult praise 54.4 (.25), peer talk 32.7 (.15), adult warn 22.8 (.11), peer boss/tease 22.0 (.10), and adult punish 18.4 (.09).

Interobserver Agreement

Analyses of individual coders would have been uninformative, because a typical coder provided relatively few observations and did not code all children. To assess reliability, we therefore examined how well two randomly selected panels of coders agreed.³ We performed univariate and multivariate within-child analyses. For the univariate analyses of each behavior rate, we correlated (over children) the rates from one panel with the rates from the other, then used the Spearman–Brown procedure to estimate reliability for both combined (McNemar, 1962). In the within-child analyses, for each child we correlated the set of four overall rates from one panel with the set of four rates from the other panel, computed the average within-child correlation, and again used the Spearman–Brown procedure. A similar approach was used for the context rates and reaction rates.

The univariate reliabilities for overall behavior rates were verbal aggression (.89), prosocial behavior (.89), withdrawal (.88), and physical aggression (.82); the mean reliability for the within-child method was .90. Univariate reliabilities for contexts were peer boss/tease (.84), adult praise (.82), peer talk (.77), adult warn (.77), adult punish (.76), and adult instruct (.69); the mean within-child reliability was .65. We examined each set of reactions separately (e.g., withdrawal to six contexts). The mean within-child reliabilities were prosocial (.91), verbal aggression (.90), physical aggression (.86), and withdrawal (.75). Univariate reliabilities for individual reactions should be interpreted with caution because some reactions (e.g., physical aggression when praised) were expected to show little variability over children. Three variables (physical and verbal aggression to praise and physical aggression to instruction) had lower variability than other measures ($SDs < .05$); these were not included in reliability summaries or interpreted further. Averaged over contexts, the mean reliabilities for reactions were verbal aggression (.64), prosocial behavior (.59), withdrawal (.56), and physical aggression (.53). Averaging over reactions, mean reliabilities were peer talk (.64), adult instruct (.63), adult praise (.60), adult punish (.58), peer boss/tease (.57), and adult warn (.50).

Results

We first examine overall behavior rates, then context rates and context-specific reactions. We used multivariate analysis of variance (MANOVA) because it does not require corrections to control Type I errors (Maxwell & Delaney, 1990). We used Hotelling T^2 to test for multivariate differences among the four Gender \times Age groups. For brevity, we report the range of significant F s ($p < .05$) rather than a list of all tests. Certain comparisons were of little interest (e.g., younger girls and older boys differ in age and gender); therefore, planned comparisons within age groups (with Bonferroni adjustments of $\alpha/2$) were used for univariate tests of gender differences. Each figure also shows tests of age differences within gender. Cohen's d is used to report effect sizes for significant main effects and planned comparisons.

Overall Behavior Rates

The patterning of overall behavior varied as a function of gender, age, and both interactively, $F_s(3, 354) = 3.03\text{--}13.18$, $p_s < .03$ (see Figure 1). Overall, girls were less physically aggressive and more prosocial than boys ($d_s = .45$ and $.34$, respectively); younger children were more physically aggressive, prosocial, and withdrawn than older ones ($d_s = .39, .29, .26$). There were multivariate differences between all Gender \times Age groups, with the smallest T^2 between older girls and boys, $T^2 = 18.22$, $F(4, 169) = 4.48$, $p < .01$. Both younger and older girls were less physically aggressive than age-matched boys ($d_s = .36, .60$). Only younger girls were reliably more prosocial than boys ($d = .37$).

Context Rates

The pattern of contexts varied as a function of gender, age, and both interactively, $F_s(5, 352) = 4.45\text{--}13.34$, $p_s < .001$ (see Figure 1). Overall, girls experienced more peer talk and less adult punish and warn than boys ($d_s = .61, .45, .26$); younger children expe-

³ Analyses are presented on data from a single summer and were restricted to children who had at least 20 hourly observations per panel and at least 2 observations of each behavior in each context per panel ($n = 101$).

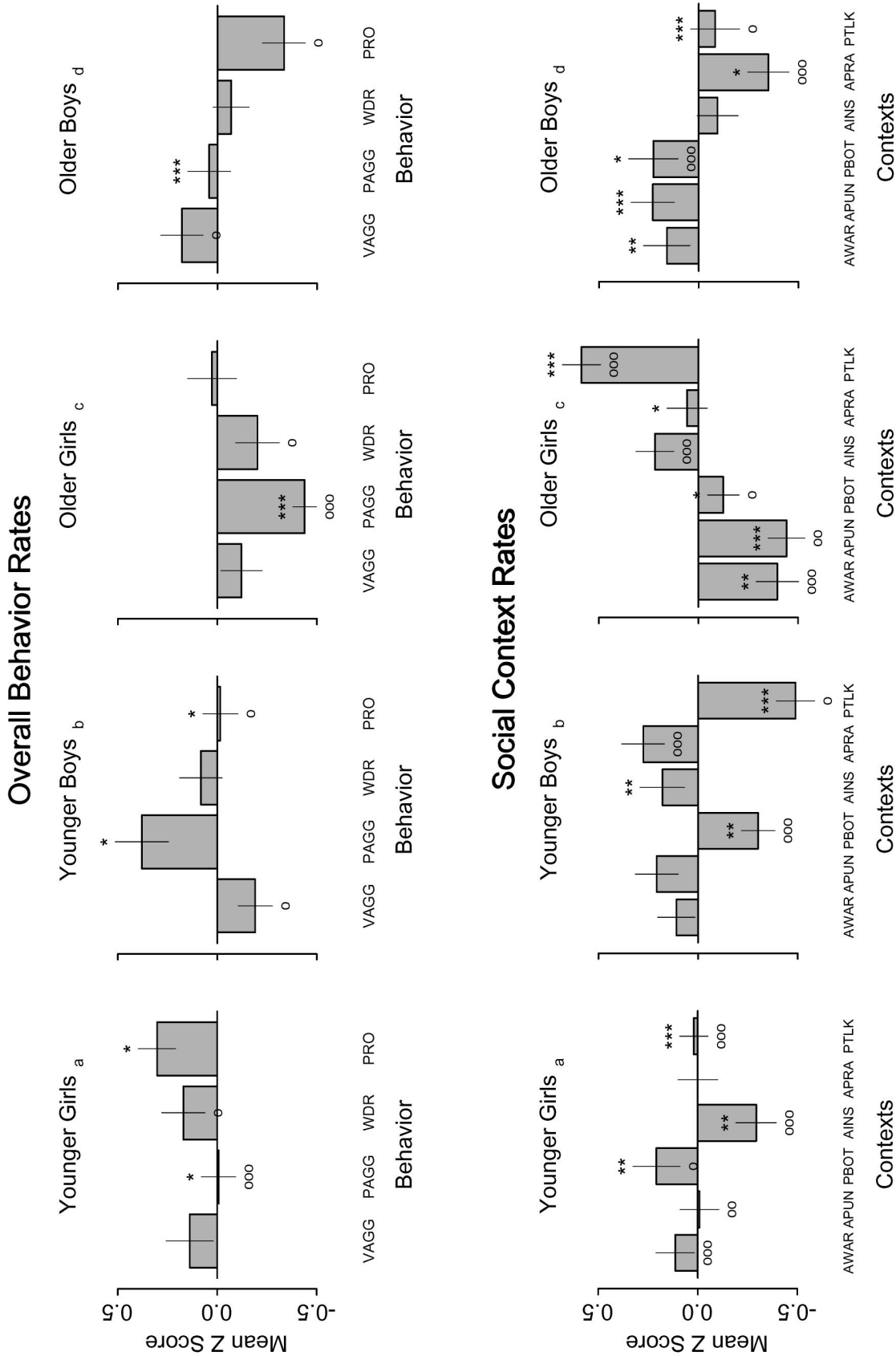


Figure 1. Mean (± 1 SEM) behavior rates (top row) and context rates (bottom row) for gender and age groups, expressed as standard deviates (z scores) from the mean for all children. Groups not sharing a subscript showed a significant multivariate difference. An asterisk indicates a significant gender comparison within age group; a circle indicates a significant age comparison within gender. Bonferroni-adjusted p values were $p < .025$, $p < .005$, and $p < .0005$ for both notations. VAGG = verbal aggression; PAGG = physical aggression; WDR = withdrawal; PRO = prosocial; AWAR = adult warm; APUN = adult punish; PBOT = peer boss/tease; AINS = adult instruct; APRA = adult praise; PTLK = peer talk.

rienced less peer talk but more adult praise and warn than older ones ($ds = .50, .29, .23$). There were multivariate differences between all Gender \times Age groups, with the smallest difference between younger girls and older boys, $T^2 = 15.57$, $F(5, 174) = 3.04$, $p < .02$.⁴ Younger girls encountered more peer talk and less adult instruct than younger boys but more peer boss/tease ($ds = .62, .46, .52$). Older girls encountered more peer talk and adult praise than older boys ($ds = .65, .43$) but less adult punish, adult warn, and peer boss/tease ($ds = .71, .54, .36$).

Reactions to Contexts

Peer boss/tease. The pattern of reactions varied as a function of gender, age, and both interactively, $F_s(3, 354) = 4.89$ – 9.56 , $ps < .002$ (see Figure 2). Girls were less likely than boys to react to peer boss/tease with physical aggression ($d = .37$) but more likely to react with withdrawn and prosocial behavior ($ds = .34, .22$); younger children were more likely than older ones to react with physical aggression and withdrawal ($ds = .42, .38$). There were multivariate differences between all Gender \times Age groups, with the smallest difference between younger girls and boys, $T^2 = 18.02$, $F(4, 181) = 4.43$, $p < .01$. Younger girls were less likely than younger boys to react with physical aggression and more likely to react with withdrawal ($ds = .39, .35$); older girls were more likely than older boys to react with prosocial and withdrawn behavior ($ds = .58, .43$) and less likely to react with physical or verbal aggression ($ds = .37, .35$).

Adult instruct. Reactions varied as a function of age and as a function of gender and age interactively, $F_s(3, 354) = 9.70$ and 4.36 , respectively, $ps < .005$ (see Figure 2). Older children were more likely than younger ones to react to adult instruction with verbal aggression ($d = .53$). All Gender \times Age groups differed multivariately, with the smallest difference between younger girls and boys, $T^2 = 11.00$, $F(4, 181) = 2.70$, $p < .04$. Younger girls were less likely than younger boys to react with physical aggression ($d = .34$); older girls were more likely than older boys to react prosocially ($d = .55$).

Adult warn. Reactions varied as a function of gender, age, and both interactively, $F_s(3, 354) = 5.34$ – 10.54 , $ps < .002$ (see Figure 3). Girls were more likely than boys to react to adult warnings with verbal aggression and less likely to react with physical aggression ($ds = .28, .22$); younger children were less likely than older ones to react with verbal aggression and more likely to react with physical aggression ($ds = .38, .22$). All groups differed multivariately except younger girls and older boys, with the smallest significant difference for older girls versus older boys, $T^2 = 15.92$, $F(4, 169) = 3.90$, $p < .01$. Younger girls were more likely than younger boys to react with verbal aggression ($d = .48$); older girls were more likely than older boys to react prosocially ($d = .42$).

Adult punish. The reactions varied as a function of gender and age, $F_s(3, 354) = 2.93$ and 10.66 , respectively, $ps < .03$ (see Figure 3). Overall, girls were more likely than boys to react to adult punishment with verbal aggression and prosocial behavior ($ds = .28, .24$); younger children were more likely than older ones to react with physical aggression and withdrawal but less likely to react with verbal aggression ($ds = .23, .21, .40$). All Gender \times Age groups differed multivariately except for older girls and older boys, with the smallest significant difference between younger girls versus older boys, $T^2 = 10.88$, $F(4, 175) = 2.67$, $p < .05$.

Planned comparisons showed that older girls were more likely than older boys to react with verbal aggression ($d = .38$).

Adult praise and peer talk. The MANOVA indicated that reactions to praise and talk varied only with age, $F_s(3, 354) = 3.99$ and 5.27 , respectively, $ps < .01$. Younger children were less likely than older ones to withdraw in response to adult praise ($d = .29$) and more likely to be physically aggressive and withdrawn in response to peer talk ($ds = .25, .27$). Although the MANOVA for gender was not significant, we report significant univariate tests for gender to compare with other results. Younger girls were less prosocial and physically aggressive to adult praise than younger boys ($ds = .42, .35$); older girls were more prosocial to peer talk than older boys ($d = .46$).

Aggregated reactions. We have noted that overall behavior rates can be affected by how often children encounter contexts and by how they react to them. An alternative approach to summarizing behavior also uses aggregation, but only over children's reactions to contexts, thus removing any effect of context rates. Specifically, we computed the average for each reaction (e.g., verbal aggression) over the six contexts (e.g., peer boss/tease, adult warn). These aggregated reactions again varied as a function of gender, age, and both interactively, $F_s(3, 354) = 5.46$ – 16.72 , $ps < .001$, but the specific findings differed from those for overall behavior rates. Overall, girls reacted with physical aggression less often than boys ($d = .35$); younger children reacted with verbal and physical aggression more often than older ones ($ds = .40, .30$). There were multivariate differences between all Gender \times Age groups, with the smallest T^2 between younger girls and older boys, $T^2 = 13.48$, $F(4, 175) = 3.31$, $p < .02$. Younger girls reacted with physical aggression less often than younger boys ($d = .45$); older girls reacted prosocially more often than older boys ($d = .57$). We compare these results with the overall behavior rates in the Discussion section.

Summary

Three sets of findings lend qualified support to our predictions. First, girls had lower overall rates of physical aggression and higher prosocial rates than boys; however, only the difference in physical aggression was reliable for both age groups. Second, we found robust differences between girls' and boys' social environments, especially for older children. Compared with boys in their age groups, older girls more often experienced peer talk and adult praise, but they less often experienced peer boss/tease, adult warn, and adult punish. Younger girls experienced more peer talk and boss/tease but less adult instruct. Third, we found localized gender differences in reactions to contexts. In both age groups, girls were less likely than boys to be aggressive in response to peer provocation and were more likely to withdraw in response to that event. However, girls were more likely than boys to be aggressive in response to direct adult control (warn or punish). Gender differences in prosocial responses depended on age but less so on context; compared with older boys, older girls were more prosocial to peer talk, peer boss/tease, adult instruct, and adult warn. Con-

⁴ Here and for any subsequent Hotelling T^2 test where it was needed to control Type I error rates, degrees of freedom were reduced by 1 because of linear dependencies among variables (see Maxwell & Delaney, 1990).

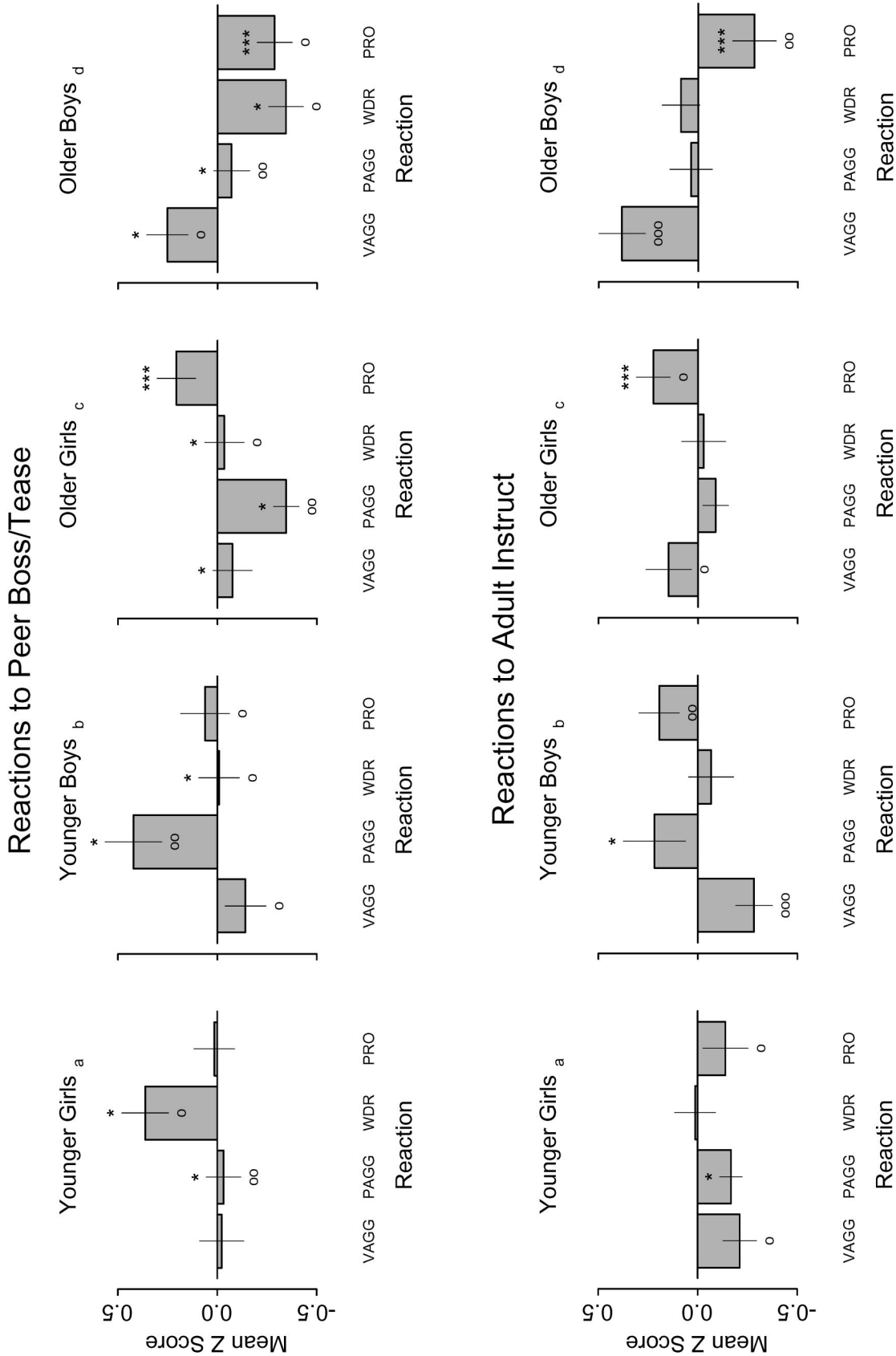


Figure 2. Mean (± 1 SEM) rates of reactions to peer boss/tease (top row) and adult instruction (bottom row) for gender and age groups, expressed as standard deviates (z scores) from the mean for all children. An asterisk indicates a significant gender comparison within age group; a circle indicates a significant age comparison within gender. Bonferroni-adjusted p values were $p < .025$, $p < .005$, and $p < .0005$ for both notations. Groups not sharing a subscript showed a significant multivariate difference. VAGG = verbal aggression; PAGG = physical aggression; WDR = withdrawal; PRO = prosocial behavior.

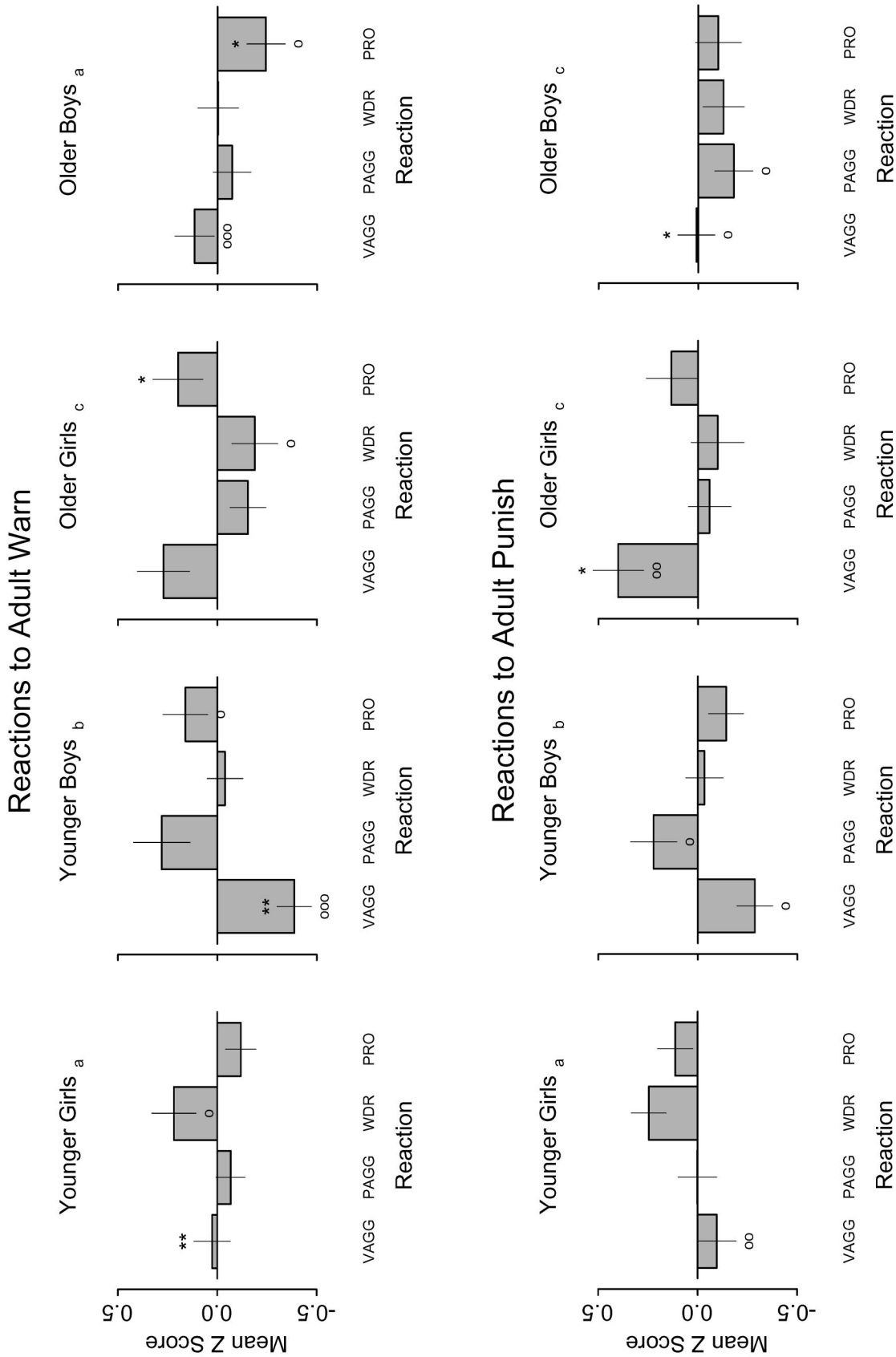


Figure 3. Mean (± 1 SEM) rates of reactions to adult warning (top row) and adult punishment (bottom row) for gender and age groups, expressed as standard deviates (z scores) from the mean for all children. An asterisk indicates a significant gender comparison within age group; a circle indicates a significant age comparison within gender. Bonferroni-adjusted p values were $p < .025$, $p < .005$, and $p < .0005$ for both notations. Groups not sharing a subscript showed a significant multivariate difference. VAGG = verbal aggression; PAGG = physical aggression; WDR = withdrawal; PRO = prosocial behavior.

sistent with our other analyses, we found that two ways of summarizing behavior—an overall act frequency measure that is affected by context rates and an aggregated reaction measure that is not—revealed distinct sets of gender and age differences.

Discussion

This research examined how a contextual model of personality can deepen our understanding of gender differences and similarities. Consistent with past work using overall behavior rates, we found differences for physical aggression and prosocial behavior but not for verbal aggression or withdrawal (Eisenberg & Fabes, 1998; Hyde, 1984; Österman et al., 1998; Rubin et al., 2002). What is notable about these findings is not only what they reveal but also what they conceal. In some cases, differences in overall behavior obscured the fact that girls and boys differed in the contexts they encountered but not in their responses to them. In other cases, similarities in overall behavior obscured the fact that girls and boys differed both in their environments and in their response patterns.

Deconstructing Differences and Similarities

Gender researchers often suggest that differences in social experiences could account for gender differences in behavior. Indeed, our results revealed robust context effects. In both peer and adult interactions, older girls' experiences were more positive than older boys', with effect sizes that were larger than those for overall behavior rates. Although it might be argued that such context effects explain away gender differences in behavior, our results reveal why such conclusions must be drawn with care. Older boys experienced a higher rate of peer provocation than older girls, but they were also more likely to be physically aggressive in response to that event. Younger children's prosocial behavior illustrates a different pattern. Younger girls, who exhibited more prosocial behavior than younger boys, were not more likely to be prosocial to any context we assessed but were more likely to encounter contexts that elicit such behavior (peer talk). Older boys' physical aggression thus was linked both to social opportunities and to heightened reactivity, whereas younger girls' prosocial behavior was linked to opportunities alone.

Our results illustrate how similarities in girls' and boys' overall behavior rates can mask other gender differences. One type of masking occurs when aggregation obscures localized gender differences. Older girls were verbally aggressive when punished by adults but not when bossed or teased by peers; older boys showed the reverse pattern. Not surprisingly, measures that aggregated over this variability showed no gender difference. Likewise, overall comparisons obscured a narrow but robust gender difference in withdrawal to peer provocation. A second type of masking occurs when context rates and reactions to contexts offset each other. Older girls were less likely than older boys to encounter peer boss/tease, adult warn, and adult instruct yet more likely to react prosocially when those contexts occurred. Because overall measures conflate reactions to events and the rate of encountering them, a net result of no difference is found.

Because context rates and reaction rates can either converge or diverge, gender comparisons will depend on how one's measure of overall behavior is affected by contexts, reactions, or both. Omnibus act frequency measures and average reaction rates are both

cross-situationally broad, yet they differ in that the former is sensitive to context rates, whereas the latter is not. Using the act frequency measure, older boys were more physically aggressive than older girls ($d = .60$), but the difference for prosocial behavior was unreliable ($d = .34$). Using aggregated reactions, the difference in physical aggression was small ($d = .17$), but a difference for prosocial behavior emerged ($d = .57$). Thus, older boys were more physically aggressive when the measure could be affected by their high rate of encountering conflict; older girls were more prosocial when the measure focused on their reactivity.

Our point is not that one way of defining overall behavior is intrinsically superior to the other but that each requires careful interpretation. Omnibus act frequency measures provide an efficient way to screen for problem (or other) behaviors in the child-environment system, regardless of their origins, but such measures do not provide explicit information about possible contextual influences. Aggregated reaction measures focus on the child's response properties, but this can be done only if relevant contexts are identified. Oversimplifying either tradeoff has risks. Overrelying on act frequency measures could reinforce the view that problem behavior is always "in" the child (see Scotti et al., 1996). Focusing exclusively on reactivity to contexts could lead one to neglect problem behaviors whose frequency stems primarily from a high rate of encountering eliciting contexts.

Gender Differences in Cycles of Cooperation and Coercion

A challenge for models of personality development (Block, 2002; Caspi, Elder, & Bem, 1988; Magnusson, 1992; Patterson, 1997) is to clarify how children's behaviors shape, and are shaped by, the behaviors of people with whom they interact. Several processes need to be considered, including how children select interaction partners, evoke responses from them, and manipulate their behavior (Buss, 1987). There is evidence that girls' and boys' behavioral styles predict their preferences for gender-segregated play (Moller & Serbin, 1996) and that aggressive children select aggressive peer groups (Cairns, Cairns, Neckerman, Gest, & Gariepy, 1988). However, when peer groups are assigned (as they were in our study), children's peer relations are likely to reflect evocation more than selection effects. In this view, older girls enjoy a mutually rewarding cycle in which their prosocial responses to peer talk reevoke and increase reciprocal behaviors from peers, and their conciliatory responses to provocation mitigate against escalating coercion. In contrast, older boys appear to be enmeshed in a cycle in which their counterattacks to provocation lead to escalating coercion from peers (see Coie et al., 1999).

As socializing or therapeutic agents, teachers, counselors, and other adults are often responsible for children's behaviors toward both peers and adults and are expected to manage behavior strategically. Adults' behaviors toward older girls (high praise, moderate instruction, low warn and punish) are likely to have multiple antecedents and to reflect both evocation and (hopefully therapeutic) "manipulation." For example, adults may use praise often to encourage older girls' positive peer relations. When adults use mild pressure (instruction), older girls respond positively, reducing the need for more direct control, but when adults do use direct control, older girls are highly reactive (verbal aggression). Older girls' responsiveness to minimal pressure and their resistance to

direct control may encourage adults to use minimal pressure to manage their behavior. High rates of peer conflict among older boys presumably require more frequent adult intervention, and boys' responses to such intervention—including their resistance to instruction but acceptance of punishments—help shape adults' greater reliance on direct control.

The finding that gender differences in reaction patterns were clearer among older children supports the view that girls and boys diverge as they adapt to different social experiences (Leaper, 2000; Maccoby, 1998). However, two developmental consistencies deserve note. First, in their peer relations, both younger and older girls were more withdrawn and less physically aggressive to peer provocation than age-matched boys. These and related results in even younger children (Fabes & Eisenberg, 1992) reveal how the distinct peer cultures of girls and boys may have their roots in "weak," or what we would call contextually narrow, initial gender differences (Maccoby, 1998). Second, in relations with adults, boys were more aggressive to adult instruction than girls, whereas girls were more aggressive to direct control. Greater male resistance to instruction also has been noted in toddlers (Kochanska & Aksan, 1995), suggesting it is a key feature of boys' interactions with adults. Our finding of girls' reactivity to discipline with mainly female counselors is consistent with evidence that girls' aggression is pronounced with female caregivers (Walsh et al., 2002) and suggests that girls may perceive such interventions as violations of female relationship norms (Mikolic, Parker, & Pruitt, 1997). More research is needed on the meaning of these different contexts to girls and boys and on how their reactions to them shape their development (e.g., Deaux & Major, 1987).

It could be argued that because children's environments reflect their personalities, it is unnecessary to disentangle contexts and reactions when studying personality. One difficulty with this view is that individuals vary in their social impact, and environments vary in their malleability. Young children influence their worlds, but their influence is limited compared with older children's or adults' (see Caspi & Roberts, 2001); some adults are permissive in their child-rearing practices, but others are more controlling (see Eisenberg & Fabes, 1998). Children's social interactions often involve complex reciprocal influences and constraints precisely because the people with whom they interact are also designing their own environments in a way that partly reflects their personalities and their attempts to alter the behaviors of others (e.g., increase prosocial behavior, decrease aggression). Moreover, children encounter some contexts and life stressors—divorce, unemployment, family relocation, school placement—for reasons that may have nothing to do with their preexisting behavior. Even in ongoing interactions, children's experiences sometimes result not from their own behaviors but rather from who they are, what they look like, or the stereotypes of people around them (Zebrowitz & Montepare, 1992). For example, we found few clear links between young girls' positive peer environments and their reactions; these girls may encounter higher rates of friendly behavior primarily because they are girls. Our claim is not that our findings point unequivocally to a particular interactional mechanism. Rather, it is that multiple contributing mechanisms usually need to be considered when interpreting gender differences or similarities in overall act trends.

Our findings should be interpreted in light of methodological strengths and weaknesses. Although our sample covered a wide

age range, cross-sectional studies are limited in their ability to identify processes that produce age differences. We believe the concept of behavioral signatures may be useful in research on personality development, but we do not know how well the signatures we found will generalize to other populations. A related question is whether our sample gives unfair advantage to context variables by restricting the range of individual differences. Because our sample included children with a range of behavior problems (Wright et al., 1999), our measures of aggression and withdrawal are less affected by restricted range than our measures of prosocial behavior; future research should examine the patterning of prosocial behaviors in normative samples. Given what is known about relational and indirect aggression, future work should investigate contexts that influence gender differences in these more subtle forms of aggression (Crick, Casas, & Nelson, 2002; Underwood, 2003). Our peer context results reflect mainly same-sex interactions among boys and girls; more research is needed on how sex of interactant affects children's behavior patterns (Archer, 1996); the sex of interactant in adult-child interactions also deserves study. Although observational methods have advantages, they are laborious; methods that use peer and adult reports are also needed (Wright & Zakriski, 2001). Other connections between this and related work should also be explored, including the ordering and organization of girls' and boys' context-specific reactions (Mikolic et al., 1997), what aggregated or context-specific reactions add to the prediction of life outcomes (Bolger & Schilling, 1991), how context-specific reactions are linked to single and multiple traits (Van Heck et al., 1994; Wright & Zakriski, 2001), and how they may be used to facilitate personality typing (Vansteelandt & Van Mechelen, 1998). Future work should also investigate individual differences within gender groups, especially in the contextual predictors of behavior, because too often such questions are overshadowed by attention to mean differences between gender groups (see Hare-Mustin & Marecek, 1988).

As Hershberger, Plomin, and Pedersen (1995) noted, "the most common measure of individual differences in personality research has undoubtedly been the total scale score obtained on an inventory or questionnaire" (p. 673). Although traits are increasingly acknowledged to be "complex patterns of behavior," a common view is that such patterns "need to be examined, interpreted, and aggregated across numerous situations, places and times, to arrive at a reliable and valid index of a personality trait" (Roberts & Caspi, 2001; p. 105). This creates a conflict for gender researchers and other developmentalists, who regularly confront these complex patterns but then conclude that they must reject the notion of personality to preserve them (Leaper, 2000; Lewis, 2001). Our point is not that aggregation necessarily leads to the mismeasure of traits or that context-specific measures always reveal personality more clearly. Rather, it is that overall behavior summaries afford multiple social interactional interpretations and cannot simply be assumed to point uniquely to "stable and enduring" nomothetic traits within the individual. Under certain conditions, aggregation can indeed filter out error or nuisance variation, leaving a purer and more refined measure of personality. Under other conditions, aggregation acts less as a filter than as a dragnet that catches a variety of effects that lie in its path, including contextual, dispositional, and dynamic interactions between the two. As we hope our findings illustrate, some gender differences in personality—especially those that are narrow and circumscribed yet potentially

important in shaping children's experiences—are revealed not by collapsing the matrix of contexts and behaviors but by preserving and scrutinizing the matrix itself.

References

- Allport, F. (1924). *Social psychology*. New York: Houghton Mifflin.
- Allport, G. W. (1937). *Personality: A psychological interpretation*. New York: Holt, Rinehart & Winston.
- Archer, J. (1996). Comparing women and men: What is being compared and why? *American Psychologist, 51*, 153–154.
- Archer, J. (2000). Sex differences in aggression between heterosexual partners: A meta-analytic review. *Psychological Bulletin, 126*, 651–680.
- Archer, J., Pearson, N. A., & Westeman, K. E. (1988). Aggressive behaviour of children aged 6–11: Gender differences and their magnitude. *British Journal of Social Psychology, 27*, 371–384.
- Barkley, R. A. (1988). Child behavior rating scales and checklists. In M. Rutter, A. H. Tuna, & I. S. Lann (Eds.), *Assessment and diagnosis in child psychopathology* (pp. 113–155). New York: Guilford Press.
- Baumeister, R. F. (1988). Should we stop studying sex differences altogether? *American Psychologist, 43*, 1092–1095.
- Block, J. (2002). *Personality as an affect-processing system: Toward an integrative theory*. Mahwah, NJ: Erlbaum.
- Bolger, N., & Schilling, E. A. (1991). Personality and the problems of everyday life: The role of neuroticism in exposure and reactivity to daily stressors. *Journal of Personality, 59*, 355–386.
- Brody, L. (1999). *Gender, emotion, and the family*. Cambridge, MA: Harvard University Press.
- Buhrmester, D., & Prager, K. (1995). Patterns and functions of self-disclosure during childhood and adolescence. In K. J. Rotenberg (Ed.), *Disclosure processes in children and adolescents* (pp. 10–56). New York: Cambridge University Press.
- Buss, D. (1987). Selection, evocation, and manipulation. *Journal of Personality and Social Psychology, 53*, 1214–1221.
- Buss, D., & Craik, K. H. (1983). The act frequency approach to personality. *Psychological Review, 90*, 105–126.
- Cairns, R. B., Cairns, B. D., Neckerman, H. J., Gest, S. D., & Garipey, J. L. (1988). Social networks and aggressive behavior: Peer support or peer rejection? *Developmental Psychology, 24*, 815–823.
- Caldera, Y. M., & Sciaraffa, M. A. (1998). Parent–toddler play with feminine toys: Are all dolls the same? *Sex Roles, 39*, 657–668.
- Caspi, A., Elder, G. H., & Bem, D. J. (1988). Moving away from the world: Life-course patterns of shy children. In S. Chess & M. Hertzog (Eds.), *Annual progress in child psychiatry and child development* (pp. 275–293). Philadelphia: Brunner/Mazel.
- Caspi, A., & Roberts, B. W. (2001). Personality development across the life course: The argument for change and continuity. *Psychological Inquiry, 12*, 49–66.
- Coie, J. D., Cillessen, A. H. N., Dodge, K. A., Hubbard, J. A., Schwartz, D., Lemerise, E. A., & Bateman, H. (1999). It takes two to fight: A test of relational factors and a method for assessing aggressive dyads. *Developmental Psychology, 35*, 1179–1188.
- Coie, J. D., & Dodge, K. A. (1998). Aggression and antisocial behavior. In N. Eisenberg (Ed.), *Handbook of child psychology* (pp. 779–862). New York: Wiley.
- Coombs, C. H. (1964). *A theory of data*. New York: Wiley.
- Crick, N. R., Casas, J. F., & Nelson, D. A. (2002). Toward a more comprehensive understanding of peer maltreatment: Studies of relational victimization. *Current Directions in Psychological Science, 11*, 98–101.
- Dawson, V. L., Zeitz, C. M., & Wright, J. C. (1989). Expert–novice differences in person perception: Evidence of experts' sensitivities to the organization of behavior. *Social Cognition, 7*, 1–30.
- Deaux, K., & Major, B. (1987). Putting gender into context. An interactive model of gender-related behavior. *Psychological Review, 94*, 369–389.
- Eagly, A. H. (1987). Reporting sex differences. *American Psychologist, 42*, 755–756.
- Eagly, A. H. (1995). The science and politics of comparing women and men. *American Psychologist, 50*, 145–158.
- Eagly, A. H., & Crowley, M. (1986). Gender and helping behavior: A meta-analytic review of the social psychological literature. *Psychological Bulletin, 100*, 283–308.
- Eisenberg, N., & Fabes, R. (1998). Prosocial development. In N. Eisenberg (Ed.), *Handbook of child psychology* (5th ed., Vol. 3, pp. 701–778). New York: Wiley.
- Fabes, R. A., & Eisenberg, N. (1992). Younger children's coping with interpersonal anger. *Child Development, 63*, 116–128.
- Feingold, A. (1994). Gender differences in personality: A meta-analysis. *Psychological Bulletin, 116*, 429–456.
- Goodman, S. H., Lahey, B. B., Fielding, B., Dulcan, M., Narrow, W., & Regier, D. (1997). Representativeness of youths with mental disorders: A preliminary population-based study. *Journal of Abnormal Psychology, 106*, 3–14.
- Gray, P. (2002). *Psychology* (4th ed.). New York: Worth.
- Gresham, F. M., & Noell, G. H. (1993). Issues in the classification of childhood psychopathology. In T. R. Kratochwill & R. J. Morris (Eds.), *Handbook of psychotherapy with children and adolescents* (pp. 108–128). Boston: Allyn & Bacon.
- Hare-Mustin, R. T., & Marecek, J. (1988). The meaning of difference: Gender theory, postmodernism, and psychology. *American Psychologist, 43*, 455–464.
- Hayes, S. T. (1995). *Accuracy of person perception in childhood*. Unpublished doctoral dissertation, American University.
- Hays, W. L. (1973). *Statistics for psychologists* (2nd ed.). New York: Holt, Reinhart & Winston.
- Helson, R., Jones, C., & Kwan, V. S. Y. (2002). Personality change over 40 years of adulthood: Hierarchical linear modeling analyses of two longitudinal samples. *Journal of Personality and Social Psychology, 83*, 752–766.
- Hershberger, S. L., Plomin, R., & Pedersen, N. L. (1995). Traits and metatraits: Their reliability, stability, and shared genetic influence. *Journal of Personality and Social Psychology, 69*, 673–685.
- Hoffenaar, P. J., & Hoeksma, J. B. (2002). The structure of oppositionality: Response dispositions and situational aspects. *Journal of Psychology and Psychiatry and Allied Health Disciplines, 43*, 375–385.
- Hyde, J. S. (1984). How large are gender differences in aggression? A developmental meta-analysis. *Developmental Psychology, 20*, 722–736.
- Johnson, J. A. (1999). Persons and situations: Distinguishing new wine from old wine in new bottles. *European Journal of Personality, 13*, 443–453.
- Kochanska, G., & Aksan, N. (1995). Mother–child mutually positive affect, the quality of child compliance to requests and prohibitions, and maternal control as correlates of early internalization. *Child Development, 66*, 236–254.
- Leaper, C. (2000). Gender, affiliation, and the interactive context of parent–child play. *Developmental Psychology, 36*, 381–393.
- Leaper, C., Anderson, K. J., & Sanders, P. (1998). Moderators of gender effects on parents' talk to their children: A meta-analysis. *Developmental Psychology, 34*, 3–27.
- Lewis, M. (2001). Issues in the study of personality. *Psychological Inquiry, 12*, 67–83.
- Lott, B., & Eagly, A. H. (1996). Research priorities: Should we continue to study gender differences? In M. R. Walsh (Ed.), *Women, men and gender: Ongoing debates* (pp. 15–31). New Haven, CT: Yale University Press.
- Maccoby, E. E. (1998). *The two sexes: Growing up apart, coming together*. Cambridge, MA: Harvard University Press.
- Maccoby, E. E. (2002). Gender and group process: A developmental perspective. *Current Directions in Psychological Science, 11*, 54–58.

- Magnusson, D. (1992). Individual development: A longitudinal perspective. *European Journal of Personality, 6*, 119–138.
- Maxwell, S. E., & Delaney, H. D. (1990). *Designing experiments and analyzing data: A model comparison perspective*. Belmont, CA: Wadsworth.
- McNemar, Q. (1962). *Psychological statistics*. New York: Wiley.
- Mikolic, J. M., Parker, J. C., & Pruitt, D. G. (1997). Escalation in response to persistent annoyance: Groups versus individuals and gender effects. *Journal of Personality and Social Psychology, 72*, 151–163.
- Miller, P. M., Danaher, D. L., & Forbes, D. (1986). Sex-related strategies for coping with interpersonal conflict in children ages 5–7. *Developmental Psychology, 22*, 543–548.
- Mischel, W., & Shoda, Y. (1995). A cognitive–affective system theory of personality: Reconceptualizing situations, dispositions, dynamics, and invariance in personality structure. *Psychological Review, 102*, 246–268.
- Moller, L. C., & Serbin, L. A. (1996). Antecedents of toddler gender segregation: Cognitive consonance, gender-typed toy preferences and behavioral compatibility. *Sex Roles, 35*, 445–460.
- Mullen, M., Snidman, N., & Kagan, J. (1993). Free-play behavior in inhibited and uninhibited children. *Infant Behavior and Development, 16*, 383–389.
- Muller, R. T., Endler, N. S., & Parker, J. D. (1990). The interaction model of anxiety assessed in two public speaking situations. *Personality and Individual Differences, 11*, 371–377.
- Österman, K., Björkqvist, K., Lagerspetz, K. M., Kaukiainen, A., Landau, S. F., Fraczek, A., & Caprara, G. V. (1998). Cross-cultural evidence of female indirect aggression. *Aggressive Behavior, 24*, 1–8.
- Patterson, G. R. (1982). *Coercive family process*. Eugene, OR: Castalia.
- Patterson, G. R. (1997). Performance models for parenting: A social interactional perspective. In J. E. Grusec & L. Kuczynski (Eds.), *Parenting and children's internalization of values: A handbook of contemporary theory* (pp. 193–226). New York: Wiley.
- Roberts, B. W., & Caspi, A. (2001). Personality development and the person–situation debate: It's déjà vu all over again. *Psychological Inquiry, 12*, 104–109.
- Rubin, K. H., Burgess, K. B., & Coplan, R. J. (2002). Social withdrawal and shyness. In P. K. Smith & C. H. Hart (Eds.), *Blackwell handbook of childhood social development* (pp. 330–352). Malden, MA: Blackwell Publishers.
- Scotti, J. R., Morris, T. L., McNeil, C. B., & Hawkins, R. P. (1996). *DSM-IV and disorders of childhood and adolescence: Can structural criteria be functional?* *Journal of Consulting and Clinical Psychology, 64*, 1177–1191.
- Shoda, Y., Mischel, W., & Wright, J. C. (1989). Intuitive interactionism in person perception: Effects of situation–behavior relations on dispositional judgments. *Journal of Personality and Social Psychology, 56*, 41–53.
- Shoda, Y., Mischel, W., & Wright, J. C. (1993). Links between personality judgments and contextualized behavior patterns: Situation–behavior profiles of personality prototypes. *Social Cognition, 11*, 399–429.
- Smith, P. K., & Boulton, M. (1990). Rough and tumble play, aggression, and dominance: Perception and behavior in children's encounters. *Human Development, 33*, 271–282.
- Underwood, M. K. (2003). *Social aggression among girls*. New York: Guilford Press.
- Van Heck, G. L., Perugini, M., Caprara, G. V., & Fröger, J. (1994). The Big Five as tendencies in situations. *Personality and Individual Differences, 16*, 715–731.
- Van Mechelen, I., & Kiers, H. A. L. (1999). Individual differences in anxiety responses to stressful situations: A three-mode component analysis model. *European Journal of Personality, 13*, 409–428.
- Vansteelandt, K. (1999). A formal model for the competency–demand hypothesis. *European Journal of Personality, 13*, 429–442.
- Vansteelandt, K., & Van Mechelen, I. (1998). Individual differences in situation–behavior profiles: A triple-typology model. *Journal of Personality and Social Psychology, 75*, 751–765.
- Walsh, M. M., Pepler, D. J., & Levene, K. (2002). A model intervention for girls with disruptive behavior problems: The Earls Court Girls Connection. *Canadian Journal of Counselling, 36*, 297–311.
- Wechsler, D. (1974). *Manual for the Wechsler Intelligence Scale for Children—Revised*. San Antonio, TX: Psychological Corporation.
- Weiten, W. (2004). *Psychology: Themes and variations* (6th ed.). Belmont, CA: Wadsworth.
- Wright, J. C., & Drinkwater, M. (1997). Rationality versus accuracy of social judgment. *Social Cognition, 15*, 245–273.
- Wright, J. C., Lindgren, K., & Zakriski, A. L. (2001). Syndromal versus contextualized personality assessment: Differentiating environmental and dispositional determinants of boys' aggression. *Journal of Personality and Social Psychology, 81*, 1176–1189.
- Wright, J. C., & Mischel, W. (1987). A conditional approach to dispositional constructs: The local predictability of social behavior. *Journal of Personality and Social Psychology, 53*, 1159–1177.
- Wright, J. C., & Zakriski, A. L. (2001). A contextual analysis of externalizing and mixed syndrome boys: When syndromal similarity obscures functional dissimilarity. *Journal of Consulting and Clinical Psychology, 69*, 457–470.
- Wright, J. C., Zakriski, A. L., & Drinkwater, M. (1999). Developmental psychopathology and the patterning of behavior and environment: Distinctive situational and behavioral signatures of internalizing, externalizing, and mixed-syndrome children. *Journal of Consulting and Clinical Psychology, 67*, 95–107.
- Zarbatany, L., McDougall, P., & Hymel, S. (2000). Gender-differentiated experience in the peer culture: Links to intimacy in preadolescence. *Social Development, 9*, 62–79.
- Zebrowitz, L. A., & Montepare, J. M. (1992). Impressions of babyfaced individuals across the life span. *Developmental Psychology, 28*, 1143–1152.

Received January 21, 2004

Revision received August 16, 2004

Accepted December 29, 2004 ■

E-Mail Notification of Your Latest Issue Online!

Would you like to know when the next issue of your favorite APA journal will be available online? This service is now available to you. Sign up at <http://watson.apa.org/notify/> and you will be notified by e-mail when issues of interest to you become available!