Children’s Use and Knowledge of Display Rules for Anger Following Hypothetical Vignettes versus Following Live Peer Interaction

Elizabeth H. Parker, Julie A. Hubbard, Sally R. Ramsden, Nicole Relyea, Karen F. Dearing and Catherine M. Smithmyer, University of Delaware, and Kelly D. Schimmel, The Pennsylvania State University

Abstract

Our primary goal was to examine the correspondence between children’s self-reported use and knowledge of display rules for anger following hypothetical vignettes versus following live peer interactions. Our secondary goal was to investigate whether children’s self-reported experience and self-reported expression of anger were related to their observed anger expression, considered an observational measure of use of display rules for anger. Participants were 274 second-grade children. Children were first interviewed about their use and knowledge of display rules for anger in game-playing situations depicted through hypothetical vignettes. Several months later, children interacted with a confederate in standardized games designed to simulate the vignettes and answered the same questions about display rules. Children’s responses were moderately related across the two contexts. However, following the live interactions, compared to the hypothetical vignettes, children reported feeling less anger, expressing less anger, intending to hide their anger more, and dissembling their anger more. In addition, there were differences in the quality and quantity of strategies for hiding anger that children generated across the two contexts. Observations of anger expression were not related to self-reports of either the experience or expression of anger.

Keywords: display rules; anger; peer relationships

An important aspect of children’s emotion knowledge is the understanding of cultural guidelines for expressing emotion in social interaction. Children as young as two to three years of age appear to use these guidelines (Cole, 1986), which have been termed ‘display rules’ (Saarni, 1979; Underwood, Coie, & Herbsman, 1992). Children report using display rules to mask anger more frequently than any other emotion, perhaps because they anticipate greater negative social consequences to result from the display of anger (Saarni & von Salisch, 1993; Underwood, 1997). Augmenting this social pres-
sure to mask anger is a consistent censure of angry behavior communicated to children at an early age in most home and school environments.

Given that the management of anger is such a critical aspect of early socialization (Fabes & Eisenberg, 1992; Scheff, 1984), having a valid method for assessing children’s use and knowledge of display rules for anger is essential. Hypothetical vignettes have been used by a number of researchers for this purpose (Gnepp & Hess, 1986; Saarni, 1979, 1988; Underwood et al., 1992). Researchers often implicitly assume that the responses that children give following hypothetical vignettes correspond to the thoughts they would have or the behaviors they would enact during live interactions.

Although the correspondence between responses in hypothetical and live contexts has not been examined for the measurement of use and knowledge of display rules for anger, researchers have assessed this correspondence for several behavioral and social-cognitive constructs. The degree of relation that they find across these two contexts depends on three factors. The first factor is whether behavior or cognition is compared across hypothetical versus live contexts. Children’s hypothetical and live responses are more related when behavior is assessed (Eisenberg, Fabes, Minore, Mathy, Hanish, & Brown, 1994; Hay, Zahn-Waxler, Cummings, & Iannotti, 1992; Mize & Ladd, 1988; Murphy & Eisenberg, 1997; Weidman & Strayhorn, 1992) rather than cognition (Hughes, Robinson, & Moore, 1991). These findings suggest that using hypothetical vignettes to assess knowledge of display rules for anger, a largely cognitive construct, may be problematic.

The second factor is whether children’s hypothetical responses are measured verbally or through enacted puppet behaviors. Children’s hypothetical versus live responses correspond more closely when enacted puppet behaviors (Eisenberg et al., 1994; Hay et al., 1992; Mize & Ladd, 1988; Murphy & Eisenberg, 1997) are used in the hypothetical context compared to verbal responses (Hughes et al., 1991; Vitaro & Pelletier, 1991). Researchers have developed innovative techniques in which children can change the facial expressions that puppets display, and these techniques provide a method for measuring use of display rules (Denham & Couchoud, 1990; Kestenbaum & Gelman, 1995). However, it is not possible to use enacted puppet behavior to measure children’s knowledge of display rules for anger, again because this construct is cognitive rather than behavioral. Thus, verbal measures remain the only appropriate means of assessing children’s knowledge of display rules for anger. This measurement approach warrants investigation, given that hypothetical verbal responses do not relate closely to live responses for those behavioral and social cognitive constructs that have been studied.

Finally, the third factor is whether children’s live responses are measured in laboratory situations versus more naturalistic settings. Children’s hypothetical and live responses correspond more closely when live responses are measured in more generalized naturalistic situations (Eisenberg et al., 1994; Hay et al., 1992; Mize & Ladd, 1988; Murphy & Eisenberg, 1997; Weidman & Strayhorn, 1992), rather than in laboratory paradigms that closely simulate the hypothetical vignettes (Hughes et al., 1991; Vitaro & Pelletier, 1991). While laboratory-based observations do not provide the same level of ecological validity as naturalistic observations, they do allow for much greater standardization of the situations in which children are observed. By placing all participants in the same social situation, laboratory standardization allows for the comparison of children’s responses to identical provocations. Using a standardized approach is particularly important when measuring display rules, because
these rules are only easily assessed in specific situations (i.e., receiving a disappointing gift). Standardization is thus the only means by which researchers can observe children in a situation that will reliably elicit the display rule of interest.

The primary goal of this study was to examine the correspondence between children’s verbal self-reported use and knowledge of display rules for anger in two situations: a) following hypothetical vignettes depicting anger-provoking peer interaction, and b) immediately following live situations designed to simulate the vignettes. We hypothesized that children’s responses to questions assessing their use and knowledge of display rules for anger would differ in these two contexts. However, competing rationales made it difficult to anticipate the direction of this difference.

On the one hand, children may report using and understanding display rules in a way that permits more anger expression in the live context as opposed to the hypothetical context. Emotional arousal during the live interaction situations may cause children to believe that a greater display of anger is appropriate, by interfering with children’s ability to access their knowledge of display rules. Because of social desirability, children also may be less likely to report that they would experience or express anger following the hypothetical vignettes, while emotional arousal in the live interaction may lessen these inhibitions.

On the other hand, children may report using and understanding display rules in a way that permits less anger expression in the live interaction context as opposed to the hypothetical context. Children may feel more motivation to regulate their anger when they are actually in the live presence of a peer, as opposed to when they are only imagining that they are interacting with a peer. Asher and his colleagues have demonstrated the importance of social goals for most children (Erdley & Asher, 1996; Renshaw & Asher, 1983; Taylor & Asher, 1984), and this desire to get along with peers may be strengthened in live as opposed to hypothetical contexts. This social motivation may result in children reporting, in the live context, that the expression of anger is less appropriate than they reported in the hypothetical context.

Our second goal was to examine whether children’s self-reports of the experience of anger were related to observations of their anger expression in the live interaction context. By relating these two constructs, we hoped to gain some information about children’s use of display rules in our anger-provoking paradigm. More specifically, if there is little correspondence between children’s self-reports about angry feelings and observations of their anger expression, one possibility is that children are using display rules to disguise their angry feelings. In a few studies, researchers have examined the relation between self-reports of emotional experience and observations of emotional expression (Eisenberg, Fabes, Miller, Fultz, Shell, Mathy, & Reno, 1989; Eisenberg, McCreath, & Ahn, 1988; Olthof & Engelberts-Vaske, 1997). However, in the paradigms used in these studies, the expression of felt emotion was considered an appropriate response, and so correspondence between self-reports of emotional experience and observations of emotion expression may be indicative of socially competent functioning. In contrast, because our anger-provoking paradigm was more likely to elicit display rule use, the lack of a connection between self-reports of anger and observations of anger expression (i.e., feeling angry but not expressing anger) may be suggestive of social competence. For this reason, we hypothesized that children’s self-reports of the experience of anger (in both the hypothetical context and the live context) would not be strongly related to observations of their anger expression.
Our third goal was to investigate whether children’s self-reports of anger expression were related to observations of their anger expression in the live interaction context. In other words, we wanted to assess the validity of children’s self-reports of anger expression, both for self-reports following the hypothetical vignettes and for self-reports following the live interactions. Casey (1993; Casey & Schlosser, 1994) is the only researcher to address the question of association between self-reports of emotional expression and direct observations of that expression. Casey (1993) placed children in a paradigm in which they received either positive or negative feedback from a peer and then observed children’s emotional expressions in response to the feedback. Afterwards, she asked children to report what facial expression they had displayed. She found that self-reports and observations of facial expression were related for positive emotions, but less so for negative emotions. Because the focus of the current study was the negative emotion of anger, we hypothesized that children’s self-reports of anger expression (in both the hypothetical context and the live context) would not be strongly related to observations of their anger expression.

Social Competence Differences

Additionally, we wanted to investigate whether the proposed relations differed for peer-rejected and aggressive children compared to other children. The responses of these children were of special interest to us because their difficulties with anger regulation (Denham, McKinley, Couchoud, & Holt, 1990; Eisenberg, Fabes, Bernzweig, Karbon, Poulin, & Hanish, 1993; Eisenberg et al., 1994; Hubbard, in press; Underwood & Hurley, 1999) may contribute to their tendency to develop conduct problems (Coie, Terry, Lenox, Lochman, & Hyman, 1995; Kupersmidt & Coie, 1990). For example, Eisenberg and her colleagues found that nonconstructive coping with anger predicted children’s disruptive behavior several years later (Eisenberg, Fabes, Murphy, Maszk, Smith, & Karbon, 1995; Eisenberg, Fabes, Shepard, Murphy, Guthrie, Jones, Friedman, Poulin, & Maszk, 1997). For these reasons, the valid measurement of anger-related constructs is especially important when studying rejected and aggressive children.

First, we hypothesized that rejected and aggressive children would display less correspondence in their self-reports between the hypothetical and live contexts than other children. In two studies assessing behavioral or social cognitive constructs, findings emerged which distinguished rejected and aggressive children from their peers when live methodologies were used, while no differences emerged when hypothetical vignettes were employed (Hughes et al., 1991; Vitaro & Pelletier, 1991). Specifically, we predicted that rejected and aggressive children would report using and understanding display rules in a way that permits less anger expression in the hypothetical context than in the live context. Their tendency toward heightened emotional arousal in live situations (Denham et al., 1990; Eisenberg et al., 1994; Fabes & Eisenberg, 1992; Hubbard, in press; Underwood & Hurley, 1999) may result in rejected and aggressive children reporting that anger expression is more acceptable in live situations than in situations where they are simply imagining feeling angry.

Second, we hypothesized that peer-rejected and aggressive children would display greater correspondence between self-reports of the experience of anger and observations of anger expression than other children. We reasoned that peer-rejected and aggressive children would be less likely to use display rules to hide their angry feel-
ings compared to their peers, and that this lesser use of display rules would be reflected in a greater correspondence between their self-reports about feeling angry and observations of their anger expression. In support of this hypothesis, other researchers have found that children who experience social competence deficits are less likely to use display rules compared to other children, across both hypothetical vignette (Jones, Abbey, & Cumberland, 1998) and live interaction methodologies (Cole, Zahn-Waxler, & Smith, 1994).

Third, we hypothesized that peer-rejected and aggressive children would display less correspondence between self-reports of anger expression and observations of anger expression than other children. In the only study to address this hypothesis, Casey and Schlosser (1994) found that externalizing children were less accurate in reporting their facial display compared to nonexternalizing children, following a peer feedback paradigm. Based on this finding, we predicted that rejected and aggressive children would be less proficient than their peers at monitoring their anger expression during the live interaction and recalling it afterwards. Similarly, we hypothesized that these children would have more difficulty than other children imagining hypothetical social situations and accurately predicting their anger expressions.

**Gender Differences**

Finally, we were interested in exploring whether there were gender differences in the proposed relations. First, we wanted to investigate the role that gender may play in the correspondence between children’s self-reported use and knowledge of display rules for anger in hypothetical versus live interaction contexts. Studies involving the comparison of children’s responses in hypothetical versus live situations have produced inconsistent findings with regard to gender. Studies have either: a) failed to reveal gender differences (Vitaro & Pelletier, 1991), b) demonstrated that different constructs are more related across contexts for boys versus girls (Eisenberg et al., 1994), c) shown that relations are stronger for boys than girls (Murphy & Eisenberg, 1997), or d) not examined the effect of gender (Hughes et al., 1991; Weidman & Strayhorn, 1992). Because of these inconsistencies, we included gender as an exploratory variable.

Second, we wanted to explore possible gender differences in the correspondence between children’s self-reports of the experience of anger and observations of their anger expression. We hypothesized that boys would show greater correspondence between self-reports of the experience of anger and observations of anger expression than girls. We reasoned that boys would be less likely to use display rules to hide their angry feelings compared to girls, and that this lesser use of display rules would be reflected in a greater correspondence between their self-reports about feeling angry and observations of their anger expression. In fact, several other investigators have found that boys use display rules less than girls, again across both hypothetical vignettes (Underwood, Coie, & Herbsman, 1992) and live interaction paradigms (Davis, 1995; Saarni, 1984).

Third, we wanted to investigate gender differences in the degree of correspondence between self-reports of anger expression and observations of anger expression. We predicted that girls would be more accurate at predicting and monitoring their angry expressions than boys, and thus, that girls would show greater correspondence between self-reports and observations of anger expression than boys. In the only other study to address this question, Casey (1993) found that girls were more accurate than boys in reporting their facial display following peer feedback.
Method
Overview

Classroom data were collected during three spring semesters in a total of 135 second-grade classrooms. A 15-minute individual interview with each child with parental permission included three parts: a) peer nominations for liking and disliking, b) peer nominations for aggression, and c) an interview-based assessment of use and knowledge of display rules for anger in two game-playing situations depicted through hypothetical vignettes.

A total of 274 children from these classrooms participated in laboratory data collection during three summers. Children participated in the laboratory phase of data collection during the summer immediately subsequent to the school year in which their classroom data were collected. The laboratory protocol lasted approximately 60 minutes per participant and consisted of two parts: a) participants interacting with a confederate in two anger-arousing, standardized, competitive games, designed to simulate the hypothetical vignettes used during classroom data collection, and b) an interview-based assessment of use and knowledge of display rules for anger in these two specific situations. The games were videotaped for later observational coding of both facial and nonverbal anger expression. These data were collected as part of a larger project on anger regulation processes in children’s peer interactions.

Participants for Classroom Data Collection

Participants were 2,052 children in 135 second-grade classrooms in public and parochial schools in the northeastern United States. These children were approximately eight years old. Ethnicity in the sample was as follows: 22% African American, 1% Asian American, 70% Caucasian, 6% Hispanic American, and 1% other. The average classroom participation rate was 69%, with a range from 40% to 93%.

We chose to study second-grade children because, between the ages of six and 11, children’s understanding of anger regulation becomes more sophisticated (Brown, Covell, & Abramovitch, 1991; McCoy & Masters, 1985; Meerum Terwogt & Olthof, 1989). Children in the second grade are in the midst of this developmental process, making them more appropriate for study than either children who have not begun to master these skills or children who have completed mastery of them.

Procedures for Classroom Data Collection

Interviewers were one Caucasian female assistant professor, seven Caucasian female graduate students, 35 Caucasian female undergraduate students, three Caucasian male undergraduate students, two Asian American female undergraduate students, and one African American male undergraduate student (a total of 49 interviewers). All interviewers were trained to administer the interview in a standardized format.

Sociometric Nominations. Interviewers obtained unlimited positive and negative sociometric nominations from participants. Children were asked to name an unlimited number of classmates whom ‘you like a lot’ and an unlimited number of classmates whom ‘you don’t like very much.’

Sociometric status was determined by calculating and standardizing numbers of positive and negative nominations received by each child across all children in each
classroom. Social preference (SP) scores were calculated as the standardized difference between liking and disliking scores, and social impact (SI) scores were calculated as the standardized sum of liking and disliking scores. Rejected children were those who received SP scores less than -1.00, ‘liked most’ scores less than 0, and ‘liked least’ scores greater than 0. Popular children were those who received SP scores greater than 1.00, ‘liked most’ scores greater than 0, and ‘liked least’ scores less than 0. Average children were those whose SP and SI scores were between -.50 and .50.

This procedure for collecting sociometric nominations and calculating social preference and social impact scores is very similar to the procedure outlined and validated by Coie, Dodge, and Coppotelli (1982). Terry and Coie (1991) found that the system developed by Coie et al. (1982) had greater temporal stability and discriminant validity than two other sociometric systems (i.e., Asher & Dodge, 1986; Newcomb & Bukowski, 1983). Our procedure deviated from this original procedure only in that children nominated an unlimited number of classmates as liked and disliked, rather than being limited to three nominations. There are two advantages to the use of unlimited nominations. First, the unlimited procedure provides a more reliable assessment of social status than the limited procedure (Terry, 2000). Second, Terry (1999) stated that reliable and valid sociometric data can be collected using the unlimited approach when as few as 40% of the children in a classroom participate. This figure stands in contrast to the 70% participation rate needed when using a limited nomination approach (Crick & Ladd, 1989).

**Peer Nominations for Aggression**

Interviewers asked participants to name an unlimited number of classmates who fit three nomination items designed to assess aggressive behavior. These items were: ‘kids in your class who start fights,’ ‘kids in your class who yell and call other kids mean names,’ and ‘kids in your class who hit and push other kids.’

The number of nominations that each child received for each of the three aggression items was standardized within classroom, these three standardized scores were summed, and the total score was restandardized within classroom. This calculation resulted in a standardized continuous score for aggression for each participant. All children with a total score equal to or greater than one standard deviation above the mean for the classroom were classified as Aggressive and all other children were classified as Nonaggressive.

The peer nomination approach to assessing aggression has been used in many previous studies (Coie & Dodge, 1983; Crick & Grotpeter, 1995; Dodge, 1980; Dodge & Frame, 1982; Perry, Perry, & Rasmussen, 1986). In terms of reliability, Crick and Grotpeter (1995) obtained a Cronbach’s alpha of .94, using these exact items. In terms of validity, peer nominations for aggression in middle childhood are concurrently related to and predict later social and psychological maladjustment (Bukowski & Newcomb, 1984; Dodge, 1983; Parker & Asher, 1987).

**Hypothetical Vignettes**

To assess self-reported use and knowledge of display rules for anger, interviewers read two hypothetical vignettes to children and asked questions about display rules following each vignette. The first vignette described a child feeling angry upon losing a
game and a prize to a peer who is playing fairly. This vignette was worded as follows: ‘Pretend you are playing a board game with another boy/girl your own age whom you have just met. If you win the game, you will win a really great prize that you just picked out. You think you might get the prize, but as you keep on playing, you realize that the other boy/girl is probably going to win. You are sure that he/she is playing fair, but you are losing anyway. You are angry.’

The second vignette described a child feeling angry when he/she loses a game and a prize to a peer who is not playing fairly. This vignette was worded as follows: ‘Pretend you are playing another board game with another boy/girl your own age whom you have just met. If you win the game, you will also win a really great prize that you just picked out. You think you might get the prize, until you notice that the other boy/girl is cheating. The more he/she cheats, the more you begin to lose. You are angry.’

We chose to study the common childhood context of competitive game-playing for two reasons. First, this context is developmentally appropriate for elementary-school-aged children, who play structured games frequently. Second, both theory (Taylor & Asher, 1984) and pilot-testing indicated that this situation is emotionally arousing for children of this age, because of their emphasis on winning and losing.

Within the context of losing a competitive game, we were interested in the situations of fair versus unfair play, because we hypothesized that these situations might produce different degrees of anger in children. However, we decided not to counterbalance the order of the ‘fair’ and ‘cheating’ situations, for either the hypothetical vignettes or the live interactions. Rather, the first vignette or interaction involved losing to a fair partner, and the second vignette or interaction involved losing to a partner who cheated. We realize that the lack of counterbalancing means that any game effects will be difficult to interpret, in that they may be the result of playing with a fair versus a cheating partner, they may be the result of losing initially versus losing repeatedly, or they may be the result of a combination of these two influences. However, given that the two games were similar except for the fair versus cheating behavior of the partner, they represent a cumulative experience that would be difficult to counterbalance.

**Assessment of Self-Reported Use and Knowledge of Display Rules for Anger**

Following each of the two vignettes, participants were asked four questions designed to assess their self-reported use and knowledge of display rules for anger. Although utilized for the first time in this study, the interview has strong face validity. First, to assess level of anger experienced, children were asked ‘How angry would you feel? Would you feel a little angry, or some angry, or a lot angry?’ This question was accompanied by a three-point Likert scale appropriate for second-grade children. Children were not given the option to report that they did not feel angry, because the purpose of the hypothetical vignettes was for them to imagine that they were angry; therefore, all subsequent questions were contingent upon this assumption.

The second question assessed the extent to which children reported that they would display versus dissemble their angry feelings facially. Participants were told, ‘Sometimes kids’ faces look the same as how they feel, and sometimes their faces look different than how they feel. I’m going to show you some faces, and I want you to point to how you would look if this happened to you. Now remember, the last question was
about how you would feel, this question is about how you would look. Would you look
happy, or not at all angry, or a little angry, or some angry, or a lot angry?’ This ques-
tion was accompanied by drawings of a face with a happy expression, a face with a
neutral expression, and three faces with increasingly angry expressions. These five
facial expressions represent a continuum of outward expression ranging from positive
to neutral to increasingly negative.

The third question assessed participants’ intention to dissemble their anger. Partici-
pants were asked, ‘Would you want the other boy/girl playing the game to find out that
you were angry or would you want to hide feeling angry so that the other boy/girl
playing the game did not find out?’

The fourth question assessed knowledge of strategies for the dissemblance of anger.
If participants responded to the third question that they would want to hide feeling
angry, the fourth question was ‘What is something that you could do so that the other
boy/girl playing the game did not find out? What is something else you could do?
(Repeat for up to five strategies).’ If participants responded to the third question that
they would want the peer to find out that they were angry, the fourth question was
‘Sometimes kids don’t want other people to find out that they feel angry. If you
did not want the other boy/girl in the game to find out that you were angry, what is
something that you could do so that he/she did not find out? What is something
else you could do? (Repeat for up to five strategies).’ Children were encouraged to try
to generate as many as five strategies; however, when they stated that they did not
know any more strategies and could not be encouraged to generate more, they were
allowed to stop. This procedure resulted in children generating different numbers of
strategies.

Children’s responses to the last question were coded as behavioral, cognitive, or
nonsense strategies. Coders were blind to the sociometric and aggression status of the
children. Reliability was estimated by having 20% of the interviews double-coded by
two assistants who were blind to which interviews were reliability trials. Examples of
strategies coded as behavioral were: leaving the situation, asking the peer to play a
different game, and physically covering his/her face (Kappa = .92). Examples of strat-
egies coded as cognitive included: pretending that the peer was not winning or cheat-
ing in the game, focusing his/her mind on a happier topic, and using positive self-talk
about his/her skill in the game (Kappa = .75). Finally, strategies coded in the nonsense
category were those that did not clearly fall in either of the first two categories
and were considered illogical responses to the question (e.g., ‘I would wet my hair’;
Kappa = .94).

**Participants for Laboratory Data Collection**

To recruit laboratory participants, we sent informational letters to parents explaining
the purpose and procedures of the study in detail. We subsequently called parents to
request the participation of their child. Parents were not informed of their child's social
status or aggression status; rather, they were simply told that we were inviting a wide
range of children to visit the laboratory. Our recruitment resulted in a subsample of
274 children from the 135 classrooms. Ethnicity in this sample was as follows: 21%
African American, 1% Asian American, 70% Caucasian, and 8% Hispanic American.
These children were selected based on sociometric status, aggressiveness, and gender.
Boys and girls were recruited to fill the following cells: Popular (not Aggressive),
Average (not Aggressive), Rejected-Aggressive, and Rejected-Nonaggressive. We
labeled this combined sociometric status/aggression variable Social Competence. We did not recruit Popular-Aggressive or Average-Aggressive children because of the scarcity of such children and because childhood aggression is most predictive of later negative outcomes when it co-occurs with peer rejection (Coie, Lochman, Terry, & Hyman, 1992; Coie et al., 1995). Exact numbers of children in each category are shown in Table 1, along with the mean social preference score and the mean continuous aggression score for each group. Overall, thirty-four percent of the children whose parents we called agreed to have their child participate in the laboratory paradigm and brought the child to the laboratory to do so. This percentage varied only slightly by Social Competence status and by gender (range of 30% to 40% across the eight cells of the design).

### Procedures for Laboratory Data Collection

**Anger-Arousing, Confederate-Based, Standardized, Competitive Game Paradigms.** Each of the two laboratory paradigms consisted of the participant and a confederate playing a competitive board game for approximately eight minutes. Both paradigms were designed to arouse a moderate degree of anger in participants. In the first game, the confederate appeared to play fairly, but the game was ‘rigged’ to insure that the participant lost. The second game was also ‘rigged’ to insure that the participant lost, but in addition, the confederate was trained to engage in blatant unfair play in a standardized manner. In order to insure that participants were invested in playing the game, a prize was offered to the winner of each game. The experimenter asked the participant to choose from five possible prizes the one that he/she would like to win. When the confederate was asked to choose a prize as well, he/she chose the same prize as the participant, as he/she was trained to do (please see Appendix A for complete details of the steps that we took to protect the welfare of participating children and confederates).

The game involved children pretending to be astronauts who bring back stars from space. As they moved around the game board, they collected and lost star tokens, which were placed on a score card; the first child to collect 25 tokens was the winner. ‘Rigging’ the game involved controlling which space the confederate and participant

---

**Table 1. Number of Participants, Mean Social Preference Scores, and Mean Continuous Aggression Scores for Each Social Competence x Gender Group**

<table>
<thead>
<tr>
<th>Social Competence x Gender Group</th>
<th>Number of Participants</th>
<th>Mean Social Preference Score</th>
<th>Mean Continuous Aggression Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Popular Males</td>
<td>43</td>
<td>1.39</td>
<td>-0.13</td>
</tr>
<tr>
<td>Popular Females</td>
<td>51</td>
<td>1.34</td>
<td>-0.47</td>
</tr>
<tr>
<td>Average Males</td>
<td>47</td>
<td>-0.02</td>
<td>0.10</td>
</tr>
<tr>
<td>Average Females</td>
<td>51</td>
<td>0.07</td>
<td>-0.39</td>
</tr>
<tr>
<td>Rejected-Aggressive Males</td>
<td>29</td>
<td>-2.19</td>
<td>2.33</td>
</tr>
<tr>
<td>Rejected-Aggressive Females</td>
<td>9</td>
<td>-1.87</td>
<td>1.65</td>
</tr>
<tr>
<td>Rejected-Nonaggressive Males</td>
<td>20</td>
<td>-1.52</td>
<td>0.04</td>
</tr>
<tr>
<td>Rejected-Nonaggressive Females</td>
<td>24</td>
<td>-1.36</td>
<td>-0.05</td>
</tr>
</tbody>
</table>
landed on for each turn. Thus, instead of movement around the board being determined by dice or spinners, a slide projector projected numbers and arrows through the one-way mirror onto a wall. The arrow told the children whose turn it was, and the number told how many spaces to move. Most children believed that the numbers were generated by a computer or by ‘magic,’ and only one child suspected rigging (this child is not included in the total N of 274).

The confederate was a same-age, same-sex child unfamiliar to the participant. We chose to use unfamiliar peers because it would not be possible to create standardized paradigms with familiar peers. Thirty-four children (19 Caucasian boys and 15 Caucasian girls) served as confederates. These children were recruited from an elementary school that did not participate in data collection. We asked teachers to nominate children whom they considered to be intelligent and outgoing. We then contacted the parents of these children and explained the child’s potential role as an ‘actor’ in the study. Those children whose parents granted permission by signing a detailed consent form were trained during after-school training sessions.

Confederates were trained in two skills. The first skill involved playing unfairly during the second game. At three points during this game, the confederate landed on the ‘best’ space of the game board, the only space that indicated the player should take three tokens. At each of these three points, the confederate was trained to take five star tokens, rather than three, in an obvious manner. Confederates were trained to cheat blatantly in order to insure that this behavior was not missed by the participants.

The second skill involved learning to behave in a neutral fashion at all times when not cheating. For example, confederates were taught not to initiate conversation with the participants and not to express obvious emotion, either positive or negative. Although most participants did not initiate much conversation with the confederate, if they did, confederates were taught to respond briefly and pleasantly, but not to prolong the verbal exchange. Although this approach was somewhat unnatural, it resulted in increased standardization. We considered other techniques to increase standardization across confederates, such as having the confederates follow a verbal script or make certain facial expressions. However, these techniques were hard for the confederates to follow and resulted in unrealistic behavior. Therefore, in order to balance any differences in confederate behavior (or appearance, for that matter) across the cells of the design, the number of games that each confederate played with participants from each of the four Social Competence groups was approximately equal.

We are aware that this paradigm may raise ethical concerns. We took a number of steps to protect the welfare of participating children and confederates (again, see Appendix A). For example, after all data were collected but before the debriefing, the participant and confederate played a third game that was rigged to insure that the participant won and received a prize. In addition, upon their arrival at the lab, participants were simply told that we wanted to learn how children played games together (parents had previously been instructed to give children the same information, and not to tell them that we were studying anger or that they would play with a child who played unfairly). However, at the conclusion of their participation, children were thoroughly debriefed. They were told that we had ‘rigged’ the first two games so that they could not win and that they did not lose those games because they were not a ‘good player.’ Furthermore, they were told that the children with whom they played were actors and that we had trained these children to cheat while playing with them, so that we could learn how children reacted when someone played unfairly.
Assessment of Knowledge and Use of Display Rules for Anger. A five-minute individual interview followed children’s participation in the fair game, and a second five-minute interview followed their participation in the cheating game. Interviewers were five Caucasian female graduate students and four Caucasian female undergraduate students, all trained to administer the interview in a standardized format. The format and questions for the live interview closely paralleled those of the hypothetical interview. Only necessary grammatical changes, such as verb tense, were made (please see Table 2 for these exact changes). Responses to the questions asking children to generate strategies for hiding anger were coded exactly as described above for the hypothetical interview. Reliability statistics were good: behavioral (kappa = .97), cognitive (kappa = .94), and nonsense (kappa = .98).

Data Reduction for Interview Variables

Hypothetical Vignette Interview Variables. We created a total of seven variables assessing self-reported use and knowledge of display rules for anger. Level of Anger Experienced reflected children’s response to the first question about how angry they would feel (range from 1 = a little angry to 3 = a lot angry). Level of Anger Expressed reflected children’s responses to the second question about how angry their face would look (−1 = happy, 0 = not at all angry, 1 = a little angry, 2 = some angry, and 3 = a lot angry).

Level of Dissemblance was created to represent the discrepancy between how angry children reported they would feel versus how angry of a facial expression they reported they would make. This variable was calculated by subtracting children’s report of how angry they would feel from their report of how angry their faces would look. These scores should be interpreted as follows: Positive numbers mean that children reported expressing more anger than they reported feeling, and as the positive numbers get higher, the discrepancy between these two reports increases. Negative numbers mean that children reported expressing less anger than they reported feeling, and as the negative numbers get lower, the discrepancy between these two reports increases. Finally, zeros mean that children reported expressing the same amount of anger that they reported feeling (range of −4 to 2).

Children’s reported intention to dissemble angry feelings was represented by a variable called Intention to Hide Anger. Children were asked if they would want to hide feeling angry from the other child or if they would want the other child to know they were angry. A response indicating the desire to hide anger was coded as zero, and a response indicating the desire to display anger was coded as one.

The final three variables were calculated from children’s responses to the fourth, open-ended question regarding strategies for dissembling anger. Number of Plausible Strategies for Hiding Anger represented the number of nonsense strategies generated by each child subtracted from the total number of strategies generated (range of zero to five). Percentage of Behavioral Strategies for Hiding Anger and Percentage of Cognitive Strategies for Hiding Anger represented the percentage of each child’s total number of strategies that were coded as behavioral and cognitive, respectively (range of 0% to 100%).

Live Interaction Interview Variables. The same seven variables were calculated from children’s responses to the live interview as were calculated from children’s responses to the hypothetical interview. A summary of the interview questions and the variables derived from these questions is provided in Table 2.
<table>
<thead>
<tr>
<th>Self-Report Variable</th>
<th>Question From Which Variable Was Derived</th>
<th>Range of Scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level of Anger Experienced</td>
<td>‘How angry would (did) you feel?’</td>
<td>1 = a little angry 2 = some angry 3 = a lot angry</td>
</tr>
<tr>
<td>Level of Anger Expressed</td>
<td>‘How would (did) your face look?’</td>
<td>−1 = happy 0 = not at all angry 1 = a little angry 2 = some angry 3 = a lot angry</td>
</tr>
<tr>
<td>Level of Dissemblance</td>
<td>Comparison of ‘How angry would (did) you feel?’ to ‘How would (did) your face look?’</td>
<td>Subtract score for Level of Anger Experienced from score for Level of Anger Expressed (range from −4 to 2)</td>
</tr>
<tr>
<td>Intention to Hide Anger</td>
<td>‘Would (Did) you want the other boy/girl playing the game to find out that you were angry or would (did) you want to hide feeling angry so that the other boy/girl playing the game did not find out?’</td>
<td>0 = hide 1 = find out</td>
</tr>
<tr>
<td>Number of Plausible Strategies for Hiding Anger</td>
<td>‘What is something that you could do so that the other boy/girl playing the game did not find out that you were angry? What is something else you could do? (What is something that you did so that the other boy/girl playing the game did not find out that you were angry? What is something else that you did? Is there anything else that you could have done)?’ Repeat for up to five strategies.</td>
<td>Range from 0 to 5 strategies</td>
</tr>
<tr>
<td>Percentage of Behavioral Strategies for Hiding Anger</td>
<td>Same as above</td>
<td>Range from 0% to 100%</td>
</tr>
<tr>
<td>Percentage of Cognitive Strategies for Hiding Anger</td>
<td>Same as above</td>
<td>Range from 0% to 100%</td>
</tr>
</tbody>
</table>

Note: Parentheses indicate the wording that was used in the live context, as opposed to the wording that was used in the hypothetical context.
Observational Coding of Anger Expression

Using the videotapes recorded during the laboratory paradigms, anger expression was measured using two coding systems previously shown to be reliable and valid in this context (Hubbard, in press). First, observers coded facial expressions of anger. Second, observers coded the display of two angry nonverbal behaviors relevant to and prevalent in the competitive game context.

Both of these coding systems utilized a cultural informant perspective, which is often taken in observational work on emotions (Gottman and Krokoff, 1989; Gross & Levenson, 1993; Underwood & Hurley, 1999). In this approach, observers’ knowledge of emotion expression is assumed to come from personal experience. An alternative approach would have been to use the most well-validated and fine-grained coding systems available for measuring emotion expression, such as the EMFACS system (Ekman & Friesen, 1975; Friesen & Ekman, 1984). However, from a practical perspective, these systems pose several problems. First, it would be difficult to get the close camera angles required to use these systems while children were playing the board game. Second, these coding systems typically only assess facial expressions of emotion, while we wanted to assess anger expression across both facial and nonverbal modalities. Third, these coding systems are costly and time-intensive. Camras, Ribordy, Hill, Martino, Spaccarelli, & Stefani (1988) presented convincing evidence of substantial agreement between coding of emotion expression conducted using the cultural informant approach and EMFACS-coding scores. As Camras et al. (1988) pointed out, fine-grained coding systems such as EMFACS may be more microanalytic than is needed for some research on emotion expression.

Procedures Used in the Training of Observers. Twenty-one observers (two Caucasian female graduate students, 17 Caucasian female undergraduate students, and two Asian American female undergraduate students) were trained in the use of the facial coding system. Thirteen observers (two Caucasian female graduate students, eight Caucasian female undergraduate students, two Caucasian male undergraduate students, and one Asian American female undergraduate student) were trained in the use of the nonverbal coding system. Each observer only coded participants’ anger expression using one of the two coding systems. Observers were blind to the hypotheses of the study and to the Social Competence status of the participants.

Reliability for the training trials was calculated by comparing observers’ data to coding completed by the second author. Observers were considered sufficiently trained when they reached a criterion of .80 (Cohen’s kappa) on three consecutive training trials.

Facial Expression Coding System

Observers coded on a second-by-second basis whether the participant displayed an angry, sad, happy, or neutral facial expression. Observers were taught to recognize angry facial expressions via global indicators such as when the participant furrowed his/her eyebrows inward or ‘set’ his/her mouth in a hard line. In addition, after observing each game, coders rated the intensity of the participant’s angry facial expressions during the game on a three-point scale (ranging from ‘a little’ to ‘some’ to ‘a lot’). Because this intensity rating covered the entire game, it was a subjective global rating, and so the points of the scale were not behaviorally grounded. When viewing the
videotapes, coders turned off the sound on the monitor, so that facial coding was not influenced by auditory content.

For reliability purposes, twenty percent of the videotapes were coded by two observers. Observers knew that reliability checks would occur throughout coding, but they were blind to which videotapes were reliability trials. The reliability index across all second-by-second facial coding was .81, and the reliability index for angry facial expressions was .73 (both Cohen’s kappa). In addition, the reliability index for the coding of the intensity of angry facial expressions was .66 (Pearson’s r). Coding systems using the cultural informant approach are best evaluated in terms of their capacity to achieve reliable results. Using this criterion, these reliability indices suggest that the cultural informant approach was adequate for the needs of this project.

This coding system resulted in three observational variables assessing anger expression: a) Frequency of Angry Facial Expressions, or the number of angry facial expressions that the participant displayed during the game, b) Mean Duration of Angry Facial Expressions, or the average number of seconds that the participant maintained an angry facial expression during the game, and c) Intensity of Angry Facial Expressions, or the rating that the observer assigned to represent the intensity of the participants’ angry facial expressions throughout the game. Because the duration of the games varied considerably (M\textsubscript{game1} = 414.24, SD\textsubscript{game1} = 72.44, M\textsubscript{game2} = 507.04, SD\textsubscript{game2} = 76.99), the frequency variables were adjusted accordingly by multiplying these scores by the mean duration for that game across all participants and dividing the resulting score by the actual duration of that game for that participant.

Nonverbal Behavior Coding System

Observers recorded all instances in which participants engaged in two nonverbal behaviors indicative of anger. Because these behaviors were of brief duration, this coding was conducted using an ‘event’ approach, in which the simple occurrence of behavior was noted, rather than its onset and offset. These were:

1) Using Game Materials Roughly: Examples included throwing game pieces or slamming the ‘man’ down on the board with force.
2) Displaying Frustration: Examples included swinging the fist, punching the fist into the hand, hitting one’s own head with the palm of the hand, and pretending to cry.

This coding system resulted in one observational variable, Frequency of Angry Nonverbal Behaviors, which represented the aggregated frequency with which the participant displayed either of the two angry nonverbal behaviors during the game. This variable was adjusted to account for differences across participants in the duration of the games in the same manner as is described above for facial expressions.

Again, reliability analyses were conducted by having twenty percent of the videotapes coded by two observers. The reliability indices for Using Game Materials Roughly and for Displaying Frustration were .85 and .78, respectively (both Cohen’s kappa).

Results

Correlational Analyses

Correlations between Children’s Self-reports about Display Rules in the Hypothetical and Live Contexts. Correlational analyses were conducted on the seven interview vari-
ables assessing children’s self-reported use and knowledge of display rules for anger. For each variable, we examined the association between children’s self-reports following the hypothetical vignettes and their self-reports following the live interaction. The statistics reported for the variable Intention to Hide Anger are actually Phi coefficients, rather than Pearson’s correlations, because this variable is dichotomous.

First, we conducted correlations across the entire sample of children. All variables except Percentage of Cognitive Strategies for Hiding Anger were significantly correlated between the two contexts. Next, we conducted separate correlations for each of the four Social Competence groups and for each of the two Genders. Finally, we used Fisher’s Z tests to test for the significance of the difference between planned pairs of correlations. First, for each variable, we compared the correlation for the average children to the correlation for each of the other three Social Competence groups. Two significant Social Competence differences emerged. For the variables Level of Anger Expressed and Intention to Hide Anger, rejected-aggressive children had stronger correlations across the two contexts than average children. Second, for each variable, we compared the correlation for the girls to the correlation for the boys. No significant Gender differences emerged for any of the seven variables (see Table 3).

Correlations between Children’s Self-reported Anger Experience and Self-reported Anger Expression versus Observed Anger Expression. We correlated the two self-report variables Level of Anger Experienced and Level of Anger Expressed with each of the four observational variables Frequency of Angry Facial Expressions, Mean Duration of Angry Facial Expressions, Intensity of Angry Facial Expressions, and Frequency of Angry Nonverbal Behaviors. We ran these correlations for children’s self-reports in both the hypothetical context and the live context.

Our first step was to conduct correlations across the entire sample of children. None of these correlations were significant. Next, we conducted separate correlations for each of the four Social Competence groups and for each of the two Genders. Only four significant correlations emerged. First, the relation between rejected-aggressive children’s self-reported experience of anger in the live interaction context and the frequency with which they were observed to display angry nonverbal behaviors was significant. Second, a similar finding emerged for rejected-nonaggressive children; the relation between their self-reported experience of anger in the live interaction context and the frequency with which they were observed to display angry nonverbal behaviors was also significant. Third, the relation between rejected-aggressive children’s self-report of anger expressed in the live interaction context and the mean duration of their observed angry facial expressions was significant. Fourth, the relation between rejected-aggressive children’s self-report of anger expressed in the live interaction context and the intensity of their observed angry facial expressions was significant (see Tables 4 and 5).

Finally, we used Fisher’s Z tests to test for the significance of the difference between planned pairs of correlations. First, in both the hypothetical and live contexts, we compared each of the correlations for the average children to the corresponding correlation for each of the other three Social Competence groups. Second, we compared the correlations for the girls to the corresponding correlation for the boys. No significant Social Competence differences or Gender differences emerged.

We considered running all correlational analyses for each Social Competence × Gender group and for the Fair and Cheating Games separately. However, we did not conduct these analyses because we were concerned about Type I error.
Table 3. Correlations Between Self-Reports About Display Rules for Anger in the Hypothetical and Live Contexts Across All Children and By Social Competence Groups and Genders

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Level of Anger Experienced</td>
<td>.27****</td>
<td>.29**</td>
<td>.19</td>
<td>.19</td>
<td>.39**</td>
<td>.33***</td>
<td>.24**</td>
</tr>
<tr>
<td>Level of Anger Expressed</td>
<td>.21***</td>
<td>.00</td>
<td>.20</td>
<td>.53***&lt;sub&gt;a&lt;/sub&gt;</td>
<td>.28</td>
<td>.16</td>
<td>.26**</td>
</tr>
<tr>
<td>Level of Dissemblance</td>
<td>.13*</td>
<td>-.06</td>
<td>.13</td>
<td>.33*</td>
<td>.30*</td>
<td>.07</td>
<td>.17*</td>
</tr>
<tr>
<td>Intention to Hide Anger</td>
<td>.36****</td>
<td>.29**</td>
<td>.36***</td>
<td>.63***&lt;sub&gt;a&lt;/sub&gt;</td>
<td>.21</td>
<td>.36***</td>
<td>.35***</td>
</tr>
<tr>
<td>Number of Plausible Strategies for Hiding Anger</td>
<td>.29****</td>
<td>.36***</td>
<td>.15</td>
<td>.18</td>
<td>.47**</td>
<td>.27**</td>
<td>.32***</td>
</tr>
<tr>
<td>Percentage of Behavioral Strategies for Hiding Anger</td>
<td>.30****</td>
<td>.27**</td>
<td>.32**</td>
<td>.35*</td>
<td>.30*</td>
<td>.22**</td>
<td>.38***</td>
</tr>
<tr>
<td>Percentage of Cognitive Strategies for Hiding Anger</td>
<td>.03</td>
<td>-.09</td>
<td>.01</td>
<td>.20</td>
<td>.19</td>
<td>-.12</td>
<td>.13</td>
</tr>
</tbody>
</table>

*Note: When the subscript <sub>a</sub> appears in the columns for popular, rejected-aggressive, or rejected-nonaggressive participants, it means that the correlation for that group differed significantly from the correlation for the average participants.

*= p < .05, **= p < .01, ***= p < .001, ****= p < .0001.
Table 4. Correlations Between the Self-Report Variable Level of Anger Experienced and Observations of Anger Expression Across All Children and By Social Competence Groups and Genders

<table>
<thead>
<tr>
<th>Observational Variable</th>
<th>Overall</th>
<th>Average</th>
<th>Popular</th>
<th>Rejected Agg.</th>
<th>Rejected Nonagg.</th>
<th>Girls</th>
<th>Boys</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Hypothetical Context</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frequency of Angry Facial Expressions</td>
<td>-.07</td>
<td>-.13</td>
<td>-.08</td>
<td>.05</td>
<td>.02</td>
<td>-.13</td>
<td>.06</td>
</tr>
<tr>
<td>Mean Duration of Angry Facial Expressions</td>
<td>-.09</td>
<td>-.14</td>
<td>-.05</td>
<td>.06</td>
<td>-.11</td>
<td>-.12</td>
<td>-.05</td>
</tr>
<tr>
<td>Intensity of Angry Facial Expressions</td>
<td>.03</td>
<td>-.02</td>
<td>.10</td>
<td>.12</td>
<td>-.13</td>
<td>.01</td>
<td>.05</td>
</tr>
<tr>
<td>Frequency of Angry Nonverbal Behaviors</td>
<td>.01</td>
<td>-.10</td>
<td>.10</td>
<td>.04</td>
<td>-.02</td>
<td>-.04</td>
<td>.07</td>
</tr>
<tr>
<td><strong>Live Context</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frequency of Angry Facial Expressions</td>
<td>-.04</td>
<td>-.11</td>
<td>-.06</td>
<td>.16</td>
<td>-.05</td>
<td>-.08</td>
<td>.03</td>
</tr>
<tr>
<td>Mean Duration of Angry Facial Expressions</td>
<td>-.07</td>
<td>-.09</td>
<td>-.16</td>
<td>.22</td>
<td>-.03</td>
<td>-.12</td>
<td>.00</td>
</tr>
<tr>
<td>Intensity of Angry Facial Expressions</td>
<td>.00</td>
<td>-.03</td>
<td>-.15</td>
<td>.28</td>
<td>-.10</td>
<td>-.13</td>
<td>.08</td>
</tr>
<tr>
<td>Frequency of Angry Nonverbal Behaviors</td>
<td>.09</td>
<td>-.14</td>
<td>-.05</td>
<td>.36*</td>
<td>.34*</td>
<td>.11</td>
<td>.07</td>
</tr>
</tbody>
</table>

* = p < .05.
Table 5. Correlations Between the Self-Report Variable Level of Anger Expressed and Observations of Anger Expression Across All Children and By Social Competence Groups and Genders

<table>
<thead>
<tr>
<th>Observational Variable</th>
<th>Overall</th>
<th>Average</th>
<th>Popular</th>
<th>Rejected Agg.</th>
<th>Rejected Nonagg.</th>
<th>Girls</th>
<th>Boys</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Hypothetical Context</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frequency of Angry Facial Expressions</td>
<td>−.06</td>
<td>−.01</td>
<td>−.11</td>
<td>−.22</td>
<td>.11</td>
<td>−.10</td>
<td>.01</td>
</tr>
<tr>
<td>Mean Duration of Angry Facial Expressions</td>
<td>−.02</td>
<td>−.03</td>
<td>−.09</td>
<td>.12</td>
<td>.11</td>
<td>−.05</td>
<td>.06</td>
</tr>
<tr>
<td>Intensity of Angry Facial Expressions</td>
<td>.05</td>
<td>.08</td>
<td>−.07</td>
<td>.10</td>
<td>.07</td>
<td>−.05</td>
<td>.13</td>
</tr>
<tr>
<td>Frequency of Angry Nonverbal Behaviors</td>
<td>−.08</td>
<td>−.07</td>
<td>−.09</td>
<td>−.18</td>
<td>.06</td>
<td>−.05</td>
<td>−.11</td>
</tr>
<tr>
<td><strong>Live Context</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frequency of Angry Facial Expressions</td>
<td>−.08</td>
<td>.06</td>
<td>−.17</td>
<td>.00</td>
<td>−.15</td>
<td>−.11</td>
<td>−.02</td>
</tr>
<tr>
<td>Mean Duration of Angry Facial Expressions</td>
<td>−.04</td>
<td>−.07</td>
<td>−.05</td>
<td>.36*</td>
<td>−.07</td>
<td>−.05</td>
<td>−.01</td>
</tr>
<tr>
<td>Intensity of Angry Facial Expressions</td>
<td>.03</td>
<td>.10</td>
<td>−.13</td>
<td>.32*</td>
<td>−.21</td>
<td>−.06</td>
<td>.09</td>
</tr>
<tr>
<td>Frequency of Angry Nonverbal Behaviors</td>
<td>.09</td>
<td>.06</td>
<td>−.01</td>
<td>.11</td>
<td>.22</td>
<td>.16</td>
<td>.00</td>
</tr>
</tbody>
</table>

* = p < .05.
ANOVA Analyses

The seven self-report interview variables served as dependent variables in ANOVA analyses. For each dependent variable except Intention to Hide Anger, a 4 (Social Competence: Average, Popular, Rejected-Aggressive, Rejected-Nonaggressive) × 2 (Gender) × 2 (Context: Hypothetical versus Live) × 2 (Game: Fair versus Cheating) Analysis of Variance (ANOVA) was performed. Social Competence and Gender served as between-subjects variables, and Context and Game were within-subjects variables. Because our specific research question concerned differences between the hypothetical and live contexts, only significant main effects or interactions for Context were reported and interpreted.

Because the self-report variable Intention to Hide Anger is dichotomous rather than continuous, these data were analyzed using multilevel logistic regression. Specifically, we used the Hierarchical Generalized Linear Model (HGLM) option in Bryk, Raudenbush, and Congdon’s (1996) HLM software. This software allows for the simultaneous analysis of between-subjects and within-subjects factors with binomial outcome variables. Again, only significant main effects or interactions for Context were reported and interpreted.

Level of Anger Experienced. For this self-report variable, there was a Context main effect, \( F(1, 266) = 324.80, p = .0001 \). Children reported that they would feel angrier following hypothetical vignettes (\( M = 2.10 \)) than they reported actually feeling after participating in the live interactions (\( M = 1.54 \)).

Level of Anger Expressed. For this self-report variable, a Context main effect was also found, \( F(1, 266) = 225.18, p = .0001 \). Following the hypothetical vignettes, children reported that they would facially express more anger (\( M = 1.24 \)) than they reported actually expressing when they were interviewed following the live interactions (\( M = .46 \)).

Level of Dissemblance. There was a significant main effect for Context for the self-report variable Level of Dissemblance, \( F(1, 266) = 4.61, p = .03 \). Although children in both contexts reported that they would/did express less anger facially than they reported they would/did feel, the discrepancy was less following the hypothetical vignettes (\( M = -.86 \)) than following the live interactions (\( M = -1.08 \)).

Intention to Hide Anger. For this self-report variable, there was a significant Context main effect (\( b = -.28, p < .01 \)) and a significant Context × Game interaction (\( b = -.18, p < .01 \)). For the fair game, children were as likely to report wanting to hide their anger in the hypothetical context (probability = .68) as in the live context (probability = .72). However, for the cheating game, children were less likely to report wanting to hide their anger in the hypothetical context (probability = .51) than in the live context (probability = .76). In addition, subgroup analyses revealed that the Context × Game interaction was significant for the average children (\( b = -.19, p < .01 \)) and for the popular children (\( b = -.29, p < .01 \)), but not for the rejected-aggressive children (\( b = .00, \text{n.s.} \)) or for the rejected-nonaggressive children (\( b = -.04, \text{n.s.} \)).
Number of Plausible Strategies for Hiding Anger. For this self-report variable, there was a significant Context main effect, $F(1, 266) = 7.94, p = .01$. Children generated more plausible strategies for hiding anger when responding based on hypothetical vignettes ($M = 2.27$) than when responding based on live interaction ($M = 2.01$). This main effect was qualified by a Gender $\times$ Context interaction, $F(1, 266) = 4.81, p = .03$. Girls generated more plausible strategies for how to hide their anger when responding based on hypothetical vignettes ($M = 2.41$) than when responding after the live interactions ($M = 1.95$), $F(1, 134) = 10.38, p = .001$. For boys, there was no difference in the number of strategies generated when responding to hypothetical vignettes ($M = 2.12$) versus when responding to live interactions ($M = 2.07$), $F(1, 138) = .13, p = .72$.

Percentage of Behavioral Strategies for Hiding Anger. There was a significant main effect for Context for this self-report variable, $F(1, 266) = 9.31, p = .01$. Children generated a higher percentage of behavioral strategies in response to the hypothetical vignettes ($M = .77$) than following the live interactions ($M = .70$).

Percentage of Cognitive Strategies for Hiding Anger. There was a significant main effect for Context for this self-report variable, $F(1, 266) = 6.38, p = .01$. Children generated a lower percentage of cognitive strategies in response to the hypothetical vignettes ($M = .06$) than following the live interactions ($M = .10$). A summary of all main effects for Context is included in Table 6. Because it required a different analysis, the variable Intention to Hide Anger is not included in this table.

| Table 6. Summary of Differences Between Children’s Self-Reports About Display Rules for Anger in the Hypothetical Context and the Live Context |
|---------------------------------|---|---|---|
| Self-Report Variable            | df  | F-value     | Mean for Hypothetical Context | Mean for Live Context |
| Level of Anger Experienced      | 1, 266 | 324.80**** | 2.10 | 1.54 |
| Level of Anger Expressed        | 1, 266 | 225.18**** | 1.24 | 0.46 |
| Level of Dissemblance           | 1, 266 | 4.61*       | −0.86 | −1.08 |
| Number of Plausible Strategies for Hiding Anger | 1, 266 | 7.94**      | 2.27 | 2.01 |
| Percentage of Behavioral Strategies for Hiding Anger | 1, 266 | 9.31**      | 0.77 | 0.70 |
| Percentage of Cognitive Strategies for Hiding Anger | 1, 266 | 6.38**      | 0.06 | 0.10 |

* = $p < .05$, ** = $p < .01$, *** = $p < .001$, **** = $p < .0001$. 
Discussion

The aim of the current paper was to address the correspondence among the following constructs: a) children’s self-reported use and knowledge of display rules for anger in a hypothetical context, b) children’s self-reported use and knowledge of display rules for anger in a live context, and c) observed anger expression, considered an observational measure of children’s use of display rules for anger. We first examined the correspondence between children’s self-reported use and knowledge of display rules for anger in two contexts: a) following hypothetical vignettes depicting anger-provoking peer interaction, and b) immediately following live situations designed to simulate the vignettes. When an ANOVA approach was used, children’s self-reports differed across the two contexts for all seven display rules variables that we measured. However, when a correlational approach was used, children’s self-reports were moderately but significantly correlated across the two contexts, for six of the seven variables that we measured (significant correlations ranged from .13 to .36). Thus, although children did not give the same responses in the two contexts, their responses were somewhat related.

What do these contrasting findings suggest about the appropriateness of using hypothetical vignettes to measure children’s use and knowledge of display rules for anger? On the one hand, if a researcher is interested in the comparison of different groups of children, then hypothetical vignettes may provide appropriate data, in that children’s responses relative to their peers remain somewhat consistent across the hypothetical and live contexts. Even in this case, however, the strength of the correlations across the two contexts is only moderate, with children’s responses following live interaction accounting for, at most, 13% of the variance in their responses to the hypothetical vignettes. On the other hand, the fact that children’s responses differed across the two contexts suggests that researchers should be somewhat cautious about considering children’s responses to hypothetical vignettes as representative of their actual or exact responses in live situations.

As a next step, we investigated whether observations of children’s actual anger expression were related to their self-reports in either the hypothetical context or the live context. More specifically, we were interested in whether children’s self-reports of how much anger they experienced or their self-reports of how much anger they expressed were related to observations of their actual anger expression. Across the entire sample, children’s reports of both the level of anger they experienced and the level of anger they expressed were unrelated to all observational measures of their anger expression, for both the hypothetical and live contexts. These findings are consistent with the work of Casey (1993), who found little relation between children’s self-reports of their negative emotional facial expressions and actual observations of these expressions.

Our original thought was that a lack of correspondence between children’s self-reported experience of anger and their observed anger expression might reflect the use of display rules for anger. However, because no measure of observed anger expression was related to self-reports of either the experience or the expression of anger, we hesitate to draw inferences about children’s use of display rules for anger. Indeed, these findings raise some concern about the validity of children’s self-report of the experience of anger or their self-report of the expression of anger. The lack of correspondence between self-reports and observations of children’s anger expression suggests that researchers may obtain more valid measures of children’s use of display rules for
anger by observing children directly, rather than relying on their self-reports. Unfortunately, however, self-report methodology is unavoidable when assessing children’s knowledge of display rules for anger, because of the cognitive and internal nature of this construct.

Given that it remains necessary to use self-report when measuring children’s knowledge of display rules for anger, it may be informative to consider the differences in self-report that emerged in the current study between the hypothetical and live contexts. Overall, our results converged to suggest that children report using and understanding display rules in a way that permits less anger expression in the live context than the hypothetical context. Following the live interactions, compared to the hypothetical vignettes, children reported feeling less anger, expressing less anger, and dissembling their anger more. Additionally, when interacting with a peer who cheated, children were more likely to report that they would want to hide their anger in the live situation than in the hypothetical one. Children may have experienced an increase in their social motivation, or desire to get along with a real, as opposed to hypothetical, peer. This motivation may have resulted in their believing that the expression of anger was less appropriate.

Furthermore, children were not able to generate as many strategies for how to hide their anger in the live context compared to the hypothetical context. Thus, although the live situation may have increased children’s intention to hide their anger, the emotional arousal of the situation may have decreased their ability to generate effective strategies for meeting this goal. Finally, the quality of strategies that children generated for hiding anger was somewhat different across the two contexts, in that children generated fewer behavioral strategies and more cognitive strategies following the live interactions than the hypothetical vignettes. Developmental studies indicate that older children generate more cognitive strategies than younger children (Altshuler & Ruble, 1989; Band & Weisz, 1988). To the extent that this finding suggests that cognitive strategies are more sophisticated than behavioral strategies, children may have generated fewer, but somewhat more advanced, strategies following the live interaction than following the hypothetical vignettes.

We are interpreting differences in children’s responses between the two contexts as resulting from the administration of hypothetical vignettes versus participation in live scenarios. However, another interpretation is possible. Children were approximately three months older when they participated in the live interactions than when they were administered the hypothetical vignettes. The context differences that emerged in children’s responses may have been due to developmental maturation. Given that context effects emerged for all seven variables that we measured, though, it seems unlikely that such a short time span could result in such consistent differences in children’s responses. One way to avoid this possible confound would have been to counterbalance the order of the ‘fair’ and ‘cheating’ situations. However, we were concerned that children would remember the live interaction and be influenced by it if they were administered the hypothetical interview at a later date.

The only effect involving Game (fair versus cheating) that emerged was for the variable Intention to Hide Anger. For the fair game, children were as likely to report wanting to hide their anger in the hypothetical context as in the live context; however, for the cheating game, children were less likely to report wanting to hide their anger in the hypothetical context than in the live context. When interpreting this effect, it is important to remember that we did not counterbalance the order of the ‘fair’ and ‘cheating’ situations. Thus, this finding may be the result of playing with a fair versus
a cheating partner, it may be the result of losing initially versus losing repeatedly, or it may be the result of a combination of these two influences.

An additional goal of our study was to investigate whether peer-rejected and aggressive children differed from their peers in terms of the relations between hypothetical versus live self-reports of display rules. More specifically, we hypothesized that peer-rejected and aggressive children would display less correspondence between self-reports in the hypothetical and live contexts than other children. No support for this hypothesis emerged. In fact, the few significant Social Competence findings were in the opposite direction of this hypothesis, and we hesitate to interpret these findings because of concerns about Type I error.

Emotional arousal in live interaction apparently does not affect cognitive processing for rejected and aggressive children more than their peers, despite findings that rejected and aggressive children experience and express more anger in live anger-provoking situations than their peers (Denham et al., 1990; Eisenberg et al., 1994; Hubbard, in press; Underwood & Hurley, 1999). Perhaps by the second grade, rejected and aggressive children have a knowledge of display rules for anger that is in line with that of their peers, across both hypothetical and live contexts. In fact, for elementary-school-aged children, Underwood has found that aggressive children do not differ from nonaggressive children in self-reported use of display rules for anger (Underwood et al., 1992) and that rejected children do not differ from their peers in self-reported use of display rules for a variety of emotions (Underwood, 1997). Thus, while our overall findings may suggest that we use some caution when assessing children’s use and knowledge of display rules for anger with hypothetical vignettes, this measurement issue does not appear to be more of a concern for rejected and aggressive children than it is for their peers.

We made two further hypotheses about rejected and aggressive children. First, we predicted that rejected and aggressive children would display more correspondence between self-reports of the experience of anger and observations of anger expression than other children. Second, we expected that these children would display less correspondence between self-reports of anger expression and observations of anger expression than other children. No support emerged for either of these hypotheses; there were no Social Competence differences for any of the correlations. Our results stand in contrast to those of Casey and Schlosser (1994), who found that externalizing children were less accurate in reporting their facial display compared to nonexternalizing children. Two discrepancies between the two studies may account for the differing results. First, Casey and Schlosser’s sample included children diagnosed with an externalizing disorder based on a structured clinical interview, while our sample included children classified as rejected and aggressive based on peer nominations. Perhaps the children in Casey and Schlosser’s sample experienced more severe behavior problems than the children in our sample, and so they may display greater deficits in the ability to accurately monitor their own emotional facial expressions. Second, Casey and Schlosser’s positive peer feedback paradigm was designed to evoke positive emotion in children, while we used a game-playing paradigm that was developed to provoke negative emotion. It is possible that ‘at-risk’ children have more difficulty accurately reporting their positive emotional displays than their negative emotional displays, perhaps because they are more apt to receive feedback from parents, teachers, and peers about the display of negative emotion.

Finally, we investigated gender differences in several relations: a) in the relations between hypothetical versus live self-reports of display rules, b) in the relations
between self-reports of the experience of anger and observations of anger expression, and c) in the relations between self-reports of the expression of anger and observations of anger expression. Across all analyses, only one gender difference emerged. Girls generated more plausible strategies for how to hide their anger when responding based on hypothetical vignettes than when responding after the live peer interactions, while boys’ ability to generate strategies did not differ across contexts. Furthermore, an examination of the cell means suggests that it was girls’ superior ability to generate plausible strategies for hiding anger in the hypothetical context that accounted for the interaction; the means for girls in the live context and boys in both contexts were similar to each other. Girls may be particularly comfortable and adept at generating strategies for hiding anger in hypothetical situations, and the fact that the majority of our interviewers were female may have increased girls’ comfort level. However, girls’ ability to generate strategies for hiding anger appears to return to the levels demonstrated by boys when they are in a live-interaction situation. One possible reason for this decrease may be that our interview was designed around the premise that children would experience some degree of anger in our paradigm. If girls experienced internalizing emotions such as sadness or disappointment rather than anger in the live context (Block, Block, & Morrison, 1981; Nolen-Hoeksema & Girgus, 1994), this discrepancy may have affected their ability to generate strategies for hiding an emotion which was different from the one that they were actually feeling.

Several limitations of the present research should be addressed. First, the interview to assess children’s self-reported use and knowledge of display rules for anger, while drawing heavily on previous literature (e.g., Ekman & Friesen, 1975; Saarni, 1979, 1988; Underwood, 1997), was utilized for the first time in the present research. Second, our findings may be specific to situations involving an unfamiliar peer, rather than a friend or acquaintance. Children may believe that it is more or less appropriate to display anger to a familiar child than to an unfamiliar child, and these differences may vary across the hypothetical and live contexts. Third, these results may also be specific to second-grade children, in that children’s ability to provide valid information using hypothetical vignettes or children’s ability to self-report validly on their emotional expressions may increase with age. An important goal for future researchers would be to examine these effects for different age groups of children. Finally, our findings cannot be generalized beyond the context of competitive game-playing. Another promising avenue for future research would be to examine children’s use and knowledge of display rules for anger across hypothetical and live contexts in other commonly experienced, anger-inducing peer situations, such as verbal provocation.

References
Display Rules in Hypothetical versus Live Contexts 553


Acknowledgements

This project was supported by grant number MH57788 from the National Institute of Mental Health. We would like to thank the confederates, interviewers, and observers who worked with us on this project. Most of all, though, we appreciate the help of the children, parents, teachers, and principals who made this project possible.

Appendix A: Steps Taken to Protect the Welfare of Participating Children and Confederates

Note: Remarks in quotes come directly from the permission letters signed by the parents of participants and confederates.

a) **Permission to Withdraw.** For participants and confederates: ‘Your child will be told that he/she can withdraw from participating at any time, and that there will be no negative consequences for choosing to withdraw or choosing not to participate.’
b) **Permission from Children Themselves.** For participants and confederates: ‘Your child’s participation is voluntary. His/her consent to participate will be obtained on the (first) day of the project, using the space at the bottom of this form.’

c) **Parental Observation.** For participants and confederates: ‘You are free to observe your child’s participation in the project from behind a one-way mirror if you choose.’

d) **Everyday Nature of Anger Arousing Paradigm.** Although children became somewhat angry, no child became aroused to the point of aggression. Thus, we believe that our paradigm subjects children to situations that are no more stressful than those which they encounter everyday.

   For participants: ‘Although we intend to make your child feel mildly angry by having him/her play with a child who sometimes does not play fairly, we strongly believe that this experience does not place your child at any additional risk, since children often interact with other children who do not play fairly.’

   For confederates: ‘Although some children may become slightly angry with your child when he/she does not play fairly, we strongly believe that this experience does not place your child at any risk, since children often interact with other children who do not play fairly.’

e) **Debriefing of Participants about Confederates’ Role.** Children were told that they helped us understand how children handle their feelings when they are angry while playing with another child. They were told that the two children they played with were actors who broke the rules and/or won the prize with all of the children. They were then given the opportunity to ask questions about the procedures in which they were involved.

   For participants: ‘After the games, it will be explained to your child that the children whom they played with were “just acting” and that they played in the same way with all the children.’

   For confederates: ‘After the games, it will be explained to the other children that your child was “just acting,” that he/she played in the same way with all of the children, and that he/she is really a very fair player.’

f) **Participants’ Awareness of Videotaping.** Children were informed before the start of the session that they would be videotaped, and they were shown the videotaping equipment. They were allowed to view their tape at the completion of the session.

g) **Third Game.** Children were given the chance to compete for a prize in a third game that was ‘rigged’ to insure that the participant won; this experience allowed the participant to experience mastery and win a desirable prize to take home.

h) **Cooperative Time.** For participants and confederates: ‘The play session will end with your child and the other children sharing a snack and having some time to play together in a cooperative, noncompetitive manner.’

i) **Survey of Parent and Child Experience.** At the completion of the laboratory data collection, participants were given a survey to take home, fill out, and mail back to the university. Both parents and children were asked to rate their experience along several dimensions and to provide comments and feedback. The response was overwhelmingly positive. Parents’ overall rating was 4.83, and children’s overall rating was 4.57 (both on a scale of 1–5). Several parents asked whether their child could return for another play session or if their child could serve as an actor in our research in the future.
j) **Confederate Training.** For confederates: ‘Your child’s participation would first involve several after-school training sessions in which he/she will be taught to act out the two roles. Great care will be taken to explain the nature and purpose of our study to the children. In particular, it will be stressed to your child that we do not condone unfair play in children in general, and that in fact the purpose of our study is to help children learn how to play together better.’