

# Early Development and Pilot Testing of a Problem-Solving Skills-Training Program for Children

**James K. Nash**

*Portland State University*

**Mark W. Fraser**

**Maeda J. Galinsky**

**Lawrence L. Kupper**

*University of North Carolina at Chapel Hill*

*Objective: This pilot study examined the impact of a prototype problem-solving skills-training program, Making Choices, on proximal outcomes in 70 6th-grade students. Method: Students received three components of Making Choices and completed pretest and posttest measures of skills on each component. Paired-sample t tests were used to assess proximal effects. Baseline measures were used to identify four subgroups of children, and differences in skill acquisition across subgroups were assessed. Results: Students displayed significantly higher scores at posttest on measures of two of three proximal skills. Nonaggressive-accepted and aggressive-accepted students displayed stronger skills at posttest, and aggressive-rejected and nonaggressive-rejected students failed to show significant gains. Conclusions: Results provided preliminary evidence of the effectiveness of Making Choices and guided refinement of the prototype program.*

**Keywords:** *skills-training groups; problem-solving; children; developmental research*

A developmental research perspective provides a framework for creating and refining interventions in social work and other professions (Rothman & Thomas, 1994). The goal of developmental research is to create “reliable, practical tools of social intervention in user-ready form” (Rothman & Thomas, p. xxv). Based on this perspective, this article describes the early

---

**Authors' Note:** Correspondence concerning this article should be addressed to James K. Nash, Graduate School of Social Work, Portland State University, P.O. Box 751, Portland, OR 97207-0751, or nashj@rri.pdx.edu. This project is supported by grants from the North Carolina State Division of Mental Health, Developmental Disabilities, and Substance Abuse Services, the Z. Smith Reynolds Foundation, and the UNC-CH Center for Injury Prevention. We thank the teachers, staff, and students who participated in this study.

Research on Social Work Practice, Vol. 13 No. 4, July 2003 432-450  
DOI: 10.1177/1049731503253360  
© 2003 Sage Publications

development and pilot testing of a classroom-based problem-solving skills-training program, called *Making Choices* (Fraser, Nash, Galinsky, & Darwin, 2000). *Making Choices* is one component of a multicomponent intervention designed to address risk factors associated with youth violence, a problem with substantial social costs (Kelley, Huizinga, Thornberry, & Loeber, 1997; Koop & Lundberg, 1992; Prothrow-Stith, 1995; Snyder, 2000; Snyder & Sickmund, 1999; U.S. Public Health Service, 2001).

The focus of the article is an analysis of the impact on targeted proximal outcomes of a prototype version of *Making Choices*, implemented in a real-world, yet relatively controlled, setting. Additionally, we explore differential program effects among subgroups of children who received the program. This involved developing and using a new measure designed to assess students' mastery of the content of discrete program components. It also involved identifying child subgroups defined by baseline measures of two keystone risk factors for chronic violence: early aggressive behavior and social rejection by peers. Establishing that a program has an expected effect on proximal outcomes is a first step in demonstrating effectiveness and is a key activity in the development of prevention programs (National Institute of Mental Health, 1998; Rothman & Thomas, 1994; Valente & Dodge, 1997).

## **BACKGROUND: EARLY PHASES OF DEVELOPMENTAL RESEARCH**

### **Conceptual Framework**

During early program development, we reviewed an extensive body of literature on the correlates of aggressive behavior in childhood. We also examined recent research on the ways children process social information and make social choices. Synthesis of the results of this literature review provided a conceptual framework to guide the design phase of the project (Comer & Fraser, 1998; Fraser, 1996; Fraser et al., 2000; Nash 1999).

In this framework, deficits in information-processing skills represent an important individual-level risk factor for aggression, social isolation, and related problems in elementary- and early middle school-age children. Examples of information-processing skills include an ability to encode and accurately interpret cues in social situations, or an ability to generate multiple potential responses for meeting one's goals. A substantial body of research indicates that children who display deficits in information-processing skills are more likely to rely on aggression to solve social problems, relative to

children with strong information-processing skills (see, e.g., Burks, Laird, & Dodge, 1999; Crick & Dodge, 1994).

Reliance on aggression to solve problems is common in early childhood, but by late elementary age, most children have learned to use socially competent problem-solving strategies, involving compromise and bargaining. Children who continue to rely on aggression are likely to experience rejection by peers with stronger social problem-solving skills. For children from a variety of racial and ethnic backgrounds and children from low- through high-income families, rejection by prosocial peers often leads to social isolation and to increased exposure to aggressive peers. Absent effective intervention, children who are aggressive and rejected are at high risk of experiencing a range of problematic short- and long-term outcomes such as academic failure and delinquency (Pettit, 1997).

### **Program Design**

Building on this framework, *Making Choices* was designed to teach information-processing skills to children (Crick & Dodge, 1994). The targeted age range included children in third to sixth grades. Consistent with a developmental research perspective, early versions of the program were reviewed rigorously by national experts and piloted in multiple settings. The resulting qualitative data were the basis for a substantial revision of the program, setting the stage for further development and pilot testing. Building on these early activities, the pilot study of *Making Choices* described in this article was the first multigroup trial of three distinguishing components of the program—units focused on encoding social cues, interpreting cues, and setting social goals.

## **METHOD**

### **Intervention**

*Making Choices* is a manualized guide to practice. It is designed to help children make friends and avoid peer rejection by teaching skills for accurately processing social information, setting social goals, and implementing a problem-solving strategy with peers. Improved social information-processing skills are thought to result in more competent social problem solving and reduced reliance on aggression in social interactions with peers and adults (Crick & Dodge, 1994).

Based on recent research on the development of peer relations and on problem solving in childhood, the *Making Choices* program manual identifies a sequence of six cognitive steps by which individuals process social information. Implied above, these include encoding social cues, interpreting social cues, developing situation-specific social goals, searching for alternative responses to meet social goals, making a decision to use a particular response, and implementing a chosen response (Crick & Dodge, 1994). Each step in the information-processing sequence directly informs one unit of the *Making Choices* program, and every unit contains a set of lessons and activities designed to teach skills related to the corresponding processing step. For example, the second lesson in the unit on encoding social cues teaches skills for recognizing a variety of verbal and nonverbal cues (e.g., facial expressions, tone of voice) in a social situation. Supplementing the social information-processing content, the *Making Choices* program opens with a unit on recognizing and regulating feelings and emotions. The manual also includes content related to group process and facilitation, to help instructors make learning sessions productive.

Instructors using the program first present basic concepts didactically. Lessons emphasize peer-oriented learning and active learning experiences such as completing role-plays, practicing skills, and discussing key ideas (Hansen, Nangle, & Meyer, 1998). The program manual specifies core concepts and skills, and practitioners are encouraged to use language and examples that are developmentally and culturally appropriate for a specific group.

### Procedures

The study used a convenience sample of 70 sixth-grade students, assigned to five regular education homerooms, at a single school in a medium-size southeastern U.S. city. All students received instruction in *Making Choices* during the homeroom period, delivered by the homeroom teacher. Prior to implementing the program, teachers received training on the social information-processing framework underlying the program and on the content, process, and format of the program.

Homeroom teachers provided instruction in *Making Choices* from early February until the end of April 1998 two times per week for 20-25 minutes per session (approximately 28 sessions). Because students had completed a curriculum on emotions in the previous semester, teachers summarized material from Unit 1 on emotions. They delivered Unit 2 (encoding social cues), Unit 3 (interpreting social cues), and Unit 4 (formulating social goals) in full. The school year ended before a full implementation could be accomplished;

however, this abbreviated implementation gave us valuable information on the effectiveness of each unit delivered, as well as clues as to what needed further development in the manual. To monitor implementation and to troubleshoot, teachers received consultation at regular team meetings during the study period.

### Measures

*Aggression and peer rejection.* Homeroom teachers completed the Carolina Children's Initiative (CCI) Checklist on all study participants as a baseline measure of students' behavior and peer status and again on completion of the pilot study. Measures from the CCI Checklist used in these analyses were (a) the baseline 24-item aggressive subscale of the Child Behavior Checklist Teacher-Report Form (CBCL-TRF) (Achenbach & McConaughy, 1987) and (b) a single item ("disliked by peers"), measured at baseline, from the Social Health Profile (SHP), a 42-item instrument designed to measure behavioral adjustment in school, social competence, and peer social status (Fast Track Project, 1997). Teachers used a Likert-type scale to rate students as being disliked by peers *almost never* (0) to *almost always* (5).

The reliability and validity of the CBCL are well documented, and this instrument is widely used in research studies and clinical practice (Achenbach & McConaughy, 1987). The SHP item "disliked by peers" has face validity as a measure of rejection by peers. This item was significantly positively correlated with pretest CBCL aggression scores ( $r = .51, p < .01$ ) and significantly negatively correlated with teachers' assessment of students' popularity ( $r = -.62, p < .01$ ), providing evidence of its construct validity. Test-retest reliability was not directly assessed. However, a comparison group of 94 students completed the CCI Checklist at pretest and posttest (approximately 3 months apart) but did not receive instruction in *Making Choices*. These students' pretest and posttest scores on the item "disliked by peers" were significantly and positively correlated ( $r = .66, p < .01$ ), providing some evidence of reliability.

*Information-processing skills.* At the beginning and end of each program unit delivered in this pilot test, students completed a unit Skill Level Activity (SLA). Developed for pilot testing *Making Choices*, the SLAs were used to assess mastery of the information-processing skills taught in the program. The SLAs followed a similar format across units and used students' responses to hypothetical social situations to measure emotional recognition and processing skills (Crick & Dodge, 1994). Teachers read a short story about a common social situation, and students were asked to imagine they

were the child in the story. Then, students answered relevant questions or identified what they might do if they were in the situation. Students completed pretest and posttest SLAs for Units 2, 3, and 4. As noted earlier, intervention teachers summarized material from Unit 1 on emotions in this study, and students did not complete SLAs for Unit 1.

The SLA for Unit 2—Encoding Social Cues—was designed to measure students' ability to encode relevant social cues. Students received 10 pictures, each depicting a social situation. For each, they were instructed to circle all the relevant cues they could identify. Examples of relevant cues included (a) a friendly (or hostile) facial expression; (b) objects, such as a baseball, jump rope, or book, indicating a peer activity; and (c) a person's body language (e.g., a fist or a wave). Scores for this SLA were a count of relevant cues circled. The first author used a scoring protocol to assess the relevance of circled cues and then assigned scores at pretest and posttest for each student.

Unit 3 of the *Making Choices* program dealt with interpreting social cues. In the SLA for this unit, teachers read aloud a series of 16 short hypothetical situations. For example,

When you are eating in the cafeteria, you see a bunch of kids talking and laughing on the other side of the room. One of them looks over at your table and says something to the others. The other kids laugh even louder.

After each story, students were asked to identify the nature of the intent of someone in the story—for instance, "Why did the kids laugh?"—by selecting one of four close-ended responses: "on purpose to be friendly," "on purpose to be mean," "accidental/mistake," and "can't tell." Scores were computed by counting the number of correct responses and ranged from 0 to 16.

The fourth unit of *Making Choices* was designed to teach children to formulate positive social goals, and the SLA sought to assess their ability to distinguish prosocial and antisocial goals using 14 hypothetical situation vignettes. For example, based on the situation presented above, students were asked, "Imagine you want to make those kids all feel bad 'cause they're laughing. Is this a *friendly (helpful)* goal, or a *not-friendly (harmful)* goal?" After hearing each vignette, students selected one of two close-ended responses to identify the nature of the goal described in the story. Scores for this SLA were calculated as the number of correct responses and ranged from 0 to 14.

*Background and demographic variables.* School administrators provided demographic and background information on students, as well as data on students' grades, achievement test scores, classroom study habits, and classroom

behaviors. Grade point average was computed using a 4-point scale (4 = A, 3 = B, etc.) for each student based on report card grades received in the semester prior to the pilot test. Four composite variables were created using teacher comments on report cards concerning students' positive and negative study habits and classroom behaviors. Comments were coded based on their content as one of the following: (a) positive classroom behaviors (e.g., "a pleasure to have in class"), (b) negative classroom behaviors (e.g., "too much unnecessary talking"), (c) positive study habits (e.g., "excellent class participation"), and (d) negative study habits (e.g., "does not take advantage of extra help"). Students' scores on these four variables reflected a count of the number of relevant comments received during the semester.

### Analysis

To assess the impact of the program on targeted proximal outcomes, a series of paired-sample *t* tests was completed using data from the entire sample of 70 students. Due to concerns that the data may not meet the assumptions of this procedure (i.e., a normally distributed continuous outcome variable), we also performed Wilcoxon matched-pairs signed-ranks tests, a nonparametric procedure, to confirm significant differences found using *t* tests (Norusis, 2000). Effect sizes for significant differences in scores at posttest, relative to pretest, were calculated by dividing the mean difference score by the posttest standard deviation. An alpha level of .05 was used for all tests.

Pretest and posttest SLA scores were available for 48 students (69%) for Unit 2, for 62 students (89%) for Unit 3, and for 45 students (64%) for Unit 4. Most of the missing scores were clustered within a single homeroom. Although this teacher provided instruction in Units 2, 3, and 4, she did not provide SLA data for the Unit 2 posttest and the Unit 4 pretest and posttest. Comparisons of students in the homeroom with missing data on Units 2 and 4 with students in the remaining intervention homerooms, using chi-square and *t* tests, revealed no significant differences on pretest CBCL aggression scores, peer status, or demographic and background variables ( $ps > .05$ ).

To explore further program impact on targeted skills, each student was assigned to a peer-status subgroup, based on scores on baseline CBCL aggression and peer status (accepted versus rejected). This resulted in four peer-status subgroups: nonaggressive-accepted; nonaggressive-rejected; aggressive-rejected; and aggressive-accepted. Demographic and background correlates of peer-status subgroup membership were identified using

chi-square and Fisher's exact tests. Next, a series of paired-sample *t* tests was run to determine whether, within each peer-status subgroup, there was a significant difference at posttest in SLA score for each unit of the program delivered in this pilot study. The nonparametric Wilcoxon matched-pairs signed-ranks test was used to confirm significant differences, and effect sizes for significant differences were calculated.

We were also interested in identifying correlates of differences in SLA scores for each unit. This analysis proceeded in two steps. First, SLA difference scores—with positive difference scores reflecting stronger skills at posttest relative to pretest—served as response variables in three stepwise multiple linear regression models with backward selection. Student-level demographic variables (sex and race/ethnicity) and background variables served as predictors in these models. Due to the small sample size, race/ethnicity was collapsed into three groups for use in regression models. This was accomplished by creating two dummy variables, one indicating African American students and a second indicating students of "Other" race/ethnicity (i.e., Hispanic, Asian, multiethnic), thus treating European American students from other than Hispanic origin as a reference group. Background variables included academically gifted special education status, grade point average, and teacher-reported positive and negative classroom behavior and study habits. The goal was to eliminate from further analyses background and demographic variables that were not important predictors of the SLA difference scores and to identify significant and marginally significant ( $p < .10$ ) student-level predictors for inclusion in multilevel regression models. Reducing the number of covariates in multilevel models was desirable due to the relatively small sample size.

Second, predictors identified as significant or marginally significant in the multiple linear regression models were included as fixed, student-level effects in three multilevel regression models predicting SLA difference scores (Bryk & Raudenbush, 1992; Goldstein, 1995). Each model included a random homeroom-level error term to account for possible within-homeroom correlation of students' SLA difference scores. This correlation was likely to exist because students were grouped within homerooms for delivery of *Making Choices* and for completion of SLAs. Fitting a single-level model (e.g., using multiple linear regression) ignores this correlation and can produce biased analysis results. The Mixed procedure in SAS statistical software was used to estimate the parameters of multilevel models (Littell, Milliken, Stroup, & Wolfinger, 1996).

## RESULTS

### Sample Characteristics

Of the 70 students who received instruction in *Making Choices*, the majority were female (59%) and European American (69%). Reflecting approximately the demographics of the school, 16% ( $n = 11$ ) of the students in the sample were African American, and 11% ( $n = 8$ ) were Asian. One student was Hispanic, and two students were identified as multiethnic. A large proportion of students (47%) were academically gifted, and the mean grade point average was 3.42 ( $SD = .50$ , range 1.0-4.0). Information on students' baseline aggression and peer status scores appears below in the section on peer-status subgroups.

### SLA Scores: All Students

Results of paired-sample  $t$  tests and nonparametric tests appear in Table 1. Because results were consistent using both procedures, results of  $t$  tests are reported here. For all students, the SLA score was significantly higher at posttest relative to pretest for two of the three units delivered in this study. For Unit 2, students had higher cue-encoding skills on completing the unit, relative to their skills prior to the unit,  $t(47) = 4.52, p < .01$ , effect size = 0.78. Similarly, results indicated an increased level of skill in distinguishing prosocial goals on completing Unit 4,  $t(44) = 3.17, p < .01$ , effect size = 0.70. However, SLA scores at posttest were not significantly different from pretest scores for Unit 3 (Interpreting Cues,  $p > .05$ ).

### Peer-Status Subgroups

To create peer-status subgroups, two variables were dichotomized. First, scores on the SHP item "disliked by classmates" permitted identification of a subgroup of rejected students. Overall, teachers rated relatively few students as being disliked by their classmates: *almost always disliked* ( $n = 1, 1.5%$ ); *very often disliked* ( $n = 0$ ); *often disliked* ( $n = 2, 2.9%$ ); *sometimes disliked* ( $n = 15, 21.4%$ ); *rarely disliked* ( $n = 11, 15.7%$ ); and *almost never disliked* ( $n = 41, 58.6%$ ). For this analysis, students described by teachers as being disliked *almost never* or *rarely* were classified as accepted ( $n = 52, 74%$ ). Rejected students were those whom teachers described as being disliked by their classmates *sometimes*, *often*, *very often*, or *almost always* ( $n = 18, 26%$ ).

Relative to their numbers in the entire sample, males ( $n = 5, 17%$ ) were less likely to fall into the rejected subgroup compared with females ( $n = 13,$

**TABLE 1: Results of Paired T and Nonparametric Tests of Skill Level Activity Scores for All Students**

	Pretest Score		Posttest Score		t(df)	z <sup>a</sup>	ES <sup>b</sup>
	M	SD	M	SD			
Unit 2 (n = 48)	48.08	13.12	58.27	13.14	4.52(47)***	3.82***	0.78
Unit 3 (n = 62)	13.97	2.23	13.09	3.15	-1.77(61)	-1.54	
Unit 4 (n = 45)	12.09	2.33	13.29	1.71	3.17(44)***	2.87***	0.70

a. z statistic using nonparametric Wilcoxon matched-pairs signed-ranks test.

b. ES = Effect size, calculated as pretest-posttest difference, divided by the posttest standard deviation.

\*\*\* $p < .01$ .

32%) although this difference was not statistically significant ( $p = .173$ ). There was a significant relationship between students' race/ethnicity and peer acceptance status as rated by teachers ( $p < .01$  using Fisher's exact test). In this sample, 9 of 11 (82%) African American students were classified as rejected, compared with 8 of 48 (17%) European American students and 1 of 8 (13%) Asian students. There was also a significant relationship between students' academically gifted special education status and peer acceptance status. Of the 33 academically gifted students, 3 (9%) were classified as rejected, whereas 14 of 36 (39%) regular education students were classified as rejected,  $\chi^2 = 8.114 (1), p < .01$ .

Students' pretest scores on the 24-item Aggressive Behavior subscale of the CBCL-TRF were used to identify aggressive and nonaggressive subgroups. Possible scores on this subscale ranged from 0 to 48. Scores in this sample ranged from 0 to 40 and were positively skewed. Scores were relatively low, indicating a low level of aggression ( $M = 5.30, SD = 8.24, Mdn = 1$ ). For this analysis, students whose scores fell above the sample mean were categorized as aggressive ( $n = 24, M = 13.83, SD = 9.17$ ), and those with scores at or below the mean made up the nonaggressive group ( $n = 46, M = 0.85, SD = 1.39$ ).

There was no significant difference between the proportion of males ( $n = 11, 38%$ ) and females ( $n = 13, 32%$ ) in this sample who were classified as aggressive. African American students were significantly more likely to be rated as aggressive ( $n = 9, 82%$ ) relative to European American ( $n = 15, 32%$ ) and Asian students ( $n = 0$ ). Based on Fisher's exact test, this distribution of proportions was unlikely to be the result of chance ( $p < .001$ ). Students who were academically gifted were significantly less likely to be classified as aggressive ( $n = 6, 18%$ ) compared with regular education students ( $n = 17, 47%$ ),  $\chi^2 = 5.293 (1), p < .05$ .

Cross-classification based on aggression status and peer acceptance status resulted in four peer-status subgroups: nonaggressive-accepted ( $n = 42$ , 60%); nonaggressive-rejected ( $n = 4$ , 6%); aggressive-accepted ( $n = 10$ , 14%); and aggressive-rejected ( $n = 14$ , 20%). Table 2 presents the demographic composition of the peer status subgroups.

As shown in Table 2, there was a significant relationship between sex and students' peer-status type ( $p < .05$  using Fisher's exact test). Differences between male and female students emerged only in the two aggressive subgroups. Female students classified as aggressive were more likely to be classified as rejected, compared with males described as aggressive. Of 13 aggressive girls, 11 (85%) were rejected. The situation was reversed for boys. Of 11 aggressive boys, only 3 (27%) were described as rejected. Thus, females comprised 79% of the aggressive-rejected subgroup, whereas males comprised 80% of the aggressive-accepted subgroup. Although it must be conditioned on the size and characteristics of the sample, these results—the first, as far as we know, to partition intervention findings on the basis of peer status, aggression, and gender—suggest that being perceived as aggressive increases risk of peer rejection for girls but not for boys.

There was also a significant relationship between students' race/ethnicity and peer-status subgroup ( $p < .01$  using Fisher's exact test). As shown in Table 2, this relationship was again most apparent for students in the aggressive subgroups. Of 9 African American students classified as aggressive, 8 (89%) were described by teachers as rejected by classmates. In contrast, of 15 European American students classified as aggressive, only 6 (40%) were rejected. Thus, being rated as aggressive carried a higher risk of being rated as rejected by peers for African American students, relative to European American students.

Peer-status subgroup was significantly related to students' special education status as well ( $p < .01$  based on Fisher's exact test). Of 6 academically gifted students classified as aggressive, only 1 (17%) was described as rejected by classmates. In contrast, 17 regular education students were classified as aggressive, of whom, 12 (71%) were also classified as rejected.

### **SLA Change Scores Within Peer-Status Subgroups**

*Nonaggressive-accepted students.* Paired-sample  $t$  tests revealed that for students in the nonaggressive-accepted (i.e., mainstream) subgroup, SLA scores for Unit 2 (encoding cues) were significantly higher at posttest relative to pretest,  $t(27) = 3.88$ ,  $p < .01$ , effect size = 0.78. Results appear in Table 3. SLA scores for Unit 4 (formulating goals) were also significantly higher at

**TABLE 2: Demographic Composition of Peer Status Subgroups**

Peer Status	Sex**		Race/Ethnicity***		
	n (%)	n (%)	n (%)	n (%)	n (%)
	Female	Male	EA	AA	A
Nonaggressive-accepted ( $n = 42$ )	26 (62)	16 (38)	31 (74)	1 (2)	7 (17)
Nonaggressive-rejected ( $n = 4$ )	2 (50)	2 (50)	2 (50)	1 (25)	1 (25)
Aggressive-accepted ( $n = 10$ )	2 (20)	8 (80)	9 (90)	1 (10)	0
Aggressive-rejected ( $n = 14$ )	11 (79)	3 (21)	6 (43)	8 (57)	0
Totals	41 (59)	29 (41)	48 (69)	11(16)	8 (11)

NOTE:  $n = 70$ . Fisher's exact test revealed a significant relationship between peer status subgroup and demographic variable at \*\* $p < .05$  and \*\*\* $p < .01$ . EA = European American; AA = African American; A = Asian; additionally, three students (two identified as multiethnic and one Hispanic student) were classified as nonaggressive-accepted.

posttest,  $t(24) = 2.80$ ,  $p < .05$ , effect size = 0.84. Posttest scores for Unit 3 (interpreting cues) were not significantly different from pretest scores for this subgroup ( $p > .05$ ). Results of nonparametric analyses were consistent with those of paired-sample  $t$  tests (see Table 3).

*Nonaggressive-rejected students.* As shown in Table 3, posttest SLA scores for Units 2, 3, and 4 were not significantly different from pretest scores for students in the nonaggressive-rejected subgroup ( $ps > .05$ ). Thus, there was no evidence of skill acquisition in these students.

*Aggressive-accepted students.* Shown in Table 3, posttest SLA scores for Unit 2 were significantly different from pretest scores for students in the aggressive-accepted subgroup,  $t(7) = 2.71$ ,  $p < .05$ , effect size = 1.24. For Unit 4, there was a marginally significant difference at posttest, relative to pretest,  $t(7) = 2.31$ ,  $p < .10$ , effect size = 2.65. No difference was found for Unit 3 for students in the aggressive-accepted subgroup. Results of nonparametric tests were consistent with results of paired-sample  $t$  tests with one exception. As shown in Table 3, nonparametric analysis yielded a significant value of the  $z$  statistic ( $p < .05$ ) for Unit 4, whereas the  $t$  statistic was marginally significant.

*Aggressive-rejected students.* As shown in Table 3, SLA posttest scores were not significantly different from pretest scores in the aggressive-rejected subgroup for Units 2, 3, and 4. Similar to their nonaggressive-rejected peers, there was no evidence of skill acquisition for these students.

**TABLE 3: Results of Paired *T* and Nonparametric Tests of Skill Level Activity Scores for Peer-Status Subgroups**

Subgroup	Pretest Score		Posttest Score		<i>t</i> ( <i>df</i> )	<i>z</i> <sup>a</sup>	ES <sup>b</sup>
	M	SD	M	SD			
Nonaggressive-accepted							
Unit 2 ( <i>n</i> = 28)	49.79	12.01	60.61	13.85	3.88 (27)***	3.43***	0.78
Unit 3 ( <i>n</i> = 36)	14.08	2.68	13.47	3.24	-0.85 (35)		
Unit 4 ( <i>n</i> = 25)	12.00	2.38	13.44	1.71	2.80 (24)**	2.47**	0.84
Nonaggressive-rejected							
Unit 2 ( <i>n</i> = 3)	48.00	23.52	58.67	8.51	0.87 (2)		
Unit 3 ( <i>n</i> = 4)	15.25	0.50	14.00	3.37	-0.78 (3)		
Unit 4 ( <i>n</i> = 3)	11.33	4.62	13.67	0.58	1.00 (2)		
Aggressive-accepted							
Unit 2 ( <i>n</i> = 8)	45.00	8.83	60.00	12.13	2.71 (7)**	1.96**	1.24
Unit 3 ( <i>n</i> = 10)	13.80	1.87	12.10	2.96	-1.63 (9)		
Unit 4 ( <i>n</i> = 8)	12.25	1.75	13.63	0.52	2.31 (7)*	2.04**	2.65
Aggressive-rejected							
Unit 2 ( <i>n</i> = 9)	45.56	17.05	49.33	10.43	0.66 (8)		
Unit 3 ( <i>n</i> = 12)	13.33	1.78	12.42	3.06	-0.85 (11)		
Unit 4 ( <i>n</i> = 9)	12.44	2.13	12.44	2.46	0.00 (8)		

a. *z* statistic based on nonparametric Wilcoxon matched-pairs signed-ranks test to confirm significant differences found in paired-sample *t* tests.

b. ES = Effect size, calculated as pretest-posttest difference, divided by posttest standard deviation.

\**p* < .06. \*\**p* < .05. \*\*\**p* < .01.

In summary, posttest scores on measures of information-processing skills for students in the nonaggressive-accepted (i.e., mainstream) subgroup were significantly higher than posttest scores for two of three units delivered in this pilot study—encoding social cues (Unit 2) and formulating prosocial goals (Unit 4). Similar results were evident for students in the aggressive-accepted subgroup. Scores for students in the remaining subgroups, comprising students classified as rejected, were not significantly higher at posttest relative to pretest for Units 2, 3, and 4.

#### Correlates of SLA Difference Scores

*Unit 2.* The hierarchical linear models predicting Unit 2 SLA difference scores revealed a significant and positive student-level effect of students' positive study habits,  $t(41) = 3.13, p < .01$ , and a marginally significant and positive effect of students' positive classroom behaviors,  $t(41) = 1.75, p < .10$ . No other student-level background or demographic variables significantly

predicted difference scores on the SLA for Unit 2. The estimated random effect associated with one particular homeroom was marginally significant and positive, indicating that students in this homeroom had difference scores on the SLA for Unit 2 that were somewhat higher (i.e., better) relative to students in the other intervention homerooms.

*Unit 3.* Status as an academically gifted student was the only student-level predictor that was associated with SLA difference scores for Unit 3,  $t(55) = 1.96, p < .10$ . The multilevel model revealed a significant and negative random effect for one homeroom,  $t(55) = -3.21, p < .01$ . This effect was associated with the same homeroom for which the random effect on the difference score for Unit 2 was significant (but positive). Thus, students in this homeroom had difference scores for the SLA on Unit 3 that were significantly lower (i.e., worse) relative to students in the other intervention homerooms.

*Unit 4.* Results of the models for Unit 4 indicated that none of the student-level demographic and background variables was significantly associated with difference scores on the SLA for this unit. One particular random homeroom-level effect was significant and positive in the multilevel model,  $t(41) = 2.10, p < .05$ . Thus, students in this homeroom had significantly higher difference scores on the Unit 4 SLA, relative to students in the other intervention homerooms. This was not the homeroom for which random effects were significant or marginally significant in the models for Units 2 and 3.

#### DISCUSSION AND APPLICATIONS TO SOCIAL WORK PRACTICE

Many students who received classroom-based instruction in the first three units of the *Making Choices* appear to have improved their problem-solving skills. Students in the nonaggressive-accepted subgroup, and students identified as aggressive but not rejected by their peers, demonstrated gains in encoding social cues and in distinguishing prosocial goals. These results are encouraging and provide support for using *Making Choices* in regular education classrooms as a tool for primary prevention, a strategy consistent with the Surgeon General's 2001 report on preventing youth violence (U.S. Public Health Service, 2001). The program may be more effective in school settings if delivered by school social workers who are knowledgeable about group process and skilled in leading prevention groups.

Of note were results indicating that (a) girls rated as aggressive were more likely to be rated as rejected by peers, relative to aggressive boys, and (b) African American students rated as aggressive were more likely to be rated as rejected, relative to aggressive European American students. These results suggest that being viewed as aggressive may increase the risk of social isolation for girls and for African American students, beyond that experienced by European American boys. Moreover, students in the aggressive-rejected subgroup failed to demonstrate gains in this study. These results highlight the importance of attending to sex differences of participants and to cultural competence (e.g., of practitioners and program content) when developing and implementing skills-training programs.

Accordingly, our revised program manual includes new material on relational aggression, a form of aggression thought to be related to peer rejection in girls (Crick, 1996). Attention to cultural competence has been a priority since the earliest work on *Making Choices*. For example, studies on social information-processing mechanisms that we reviewed used samples that included large numbers of African American children. To increase the cultural competence of the program, there has been greater emphasis on collaborating with organizations such as church-based after-school programs in African American communities to develop and carry out pilot tests of *Making Choices*. Feedback from practitioners and others involved in these pilot tests has been incorporated into program content. Additionally, subsequent studies of *Making Choices*—some of which are currently under way—have included larger samples of African American students and a greater number of African American practitioners.

Our results are consistent with research suggesting that sociocultural risks affect African American children more than European American children (Deater-Deckard, Dodge, Bates, & Pettit, 1998). Deater-Decker and colleagues (1998) argue that risk among all children is cumulative. African American children often have the highest levels of cumulative risk, and their risk is principally sociocultural. Given the cumulative risk model for African American children, a multicomponent and culturally sensitive intervention is warranted. We do not think that a child-focused skills-training program alone is an adequate intervention for high-risk children. In fact, we believe that the higher the cumulative risk, the greater the need for a multielement intervention. Examples include working with families to meet basic needs and improve parenting skills (see, e.g., Fraser, Abell, Galinsky, Schopler, & Hodges, 2001) and building a greater sense of community and belonging in neighborhoods and schools, in addition to building skills at the individual level (Fraser, Richman, & Galinsky, 1999).

As a result of the findings in this study, we developed a multiple-impact, risk-focused, family-centered intervention to address sociocultural risks (related to poverty and acculturation stress) and family processes that are associated with the early onset of antisocial, aggressive behavior (Loeber, Farrington, Stouthamer-Loeber, & Van Kammen, 1998). Although findings from this study are encouraging, they indicate—as we have argued elsewhere—that multielement interventions, which address a range of risk factors across the family, school, and neighborhood domains, are needed to produce positive findings among high-risk children (Fraser & Galinsky, 1997). *Making Choices* should be viewed as but one intervention among an array of promising interventions that might be utilized to affect risk factors associated with youth violence or other social problems.

Although intriguing, limitations of the study should be kept in mind when interpreting these results. This was a small, nonprobability sample, and teachers (all of whom were European American)—not children—rated aggression and peer status. The sample we used is probably not typical of middle schools in the United States. Almost half of the students were classified as academically gifted, and the mean grade point average was 3.4. Future studies should be conducted with children who are more typical of the population of middle school students. It will also be important in future studies to include other measures of aggressive behavior and peer status (e.g., self-reports, parent reports, reports from other teachers) and to continue to examine possible interactions of sex, race/ethnicity, aggressive behavior, and peer status.

Students who may be at greatest risk of chronic aggression, due to peer rejection—especially coupled with early reliance on aggression—failed to demonstrate significant gains. However, students in the aggressive-rejected and nonaggressive-rejected subgroups did not display weaker skills at posttest, relative to pretest. Thus, it appears that the intervention did no harm. Nevertheless, differential effects across subgroups highlight a need to examine more closely how to increase the impact of the program for rejected children. Multilevel models suggest that teachers were differentially effective in delivering the program. In particular, it appears that the composition of groups and the integration of rejected children into classroom activities must be more deliberately addressed.

From a developmental research perspective, this pilot study provided information to guide further development and testing of *Making Choices*. Feedback from teachers who implemented the program indicated that, overall, the *Making Choices* program manual is user-friendly and targets skills that are important—and often lacking—for most children. The study

demonstrated that the program can be used in classrooms with a relatively large number of children. Results indicating that students failed to demonstrate gains on skills covered in Unit 3 of the program—interpreting social cues—alerted us to the need to pay special attention to this unit as we revised *Making Choices*. Changes to Unit 3 included adding new activities that focus on links between values and beliefs and cue interpretation and screening activities for cultural appropriateness.

The pilot study also revealed that using SLAs can be a feasible approach to measuring proximal targeted outcomes in students who receive instruction in *Making Choices*. The SLAs were under development in this study. They used students' responses to questions about hypothetical social situations to assess information-processing skills. This approach has been used extensively in prior research but usually via structured interviews with individual respondents (Crick & Dodge, 1994). Group-based administration of these measures represented a departure in format from prior research.

Finally, findings of significant random effects of homerooms in each of the models predicting SLA difference scores draw attention to the possibility that classroom-level characteristics influenced program impact. Maintaining treatment integrity across settings—including classrooms and schools—is a key challenge, not only in developmental research but in the delivery of all prevention services (Valente & Dodge, 1997). Teaching styles, aggregated student characteristics, physical space, and time are among the classroom-level characteristics that may influence the effects of a program such as *Making Choices*. Our findings demonstrate the importance of identifying, measuring, and, whenever possible, controlling the range of potentially influential classroom-level characteristics in intervention research.

To sum up, this pilot test suggests that *Making Choices* produces positive outcomes for some, but not all, children. Use of unit-specific SLAs to examine proximal outcomes within subgroups of children defined by keystone risk factors enabled us to assess program effects with some precision and guided further development of the program. In the context of significant contextual effects and findings that vary by gender and race/ethnicity, our current work focuses on strengthening the program and developing a complementary family-centered intervention called *Strong Families* (Fraser et al., 2001).

## REFERENCES

- Achenbach, T., & McConaughy, S. (1987). *Empirically based assessment of child and adolescent psychopathology: Practical applications*. Newbury Park, CA: Sage.
- Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical linear models*. Newbury Park, CA: Sage.

- Burks, V. S., Laird, R. D., & Dodge, K. A. (1999). Knowledge structures, social information processing, and children's aggressive behavior. *Social Development, 8*, 220-236.
- Comer, E. W., & Fraser, M. W. (1998). Evaluation of six family support programs: Are they effective? *Families in Society, 72*, 134-148.
- Crick, N. R. (1996). The role of overt aggression, relational aggression, and prosocial behavior in the prediction of children's future social adjustment. *Child Development, 67*, 2317-2327.
- Crick, N. R., & Dodge, K. A. (1994). A review and reformulation of social-information processing mechanisms in children's social adjustment. *Psychological Bulletin, 115*, 74-101.
- Deater-Deckard, K., Dodge, K. A., Bates, J. E., & Pettit, G. E. (1998). Multiple risk factors in the development of externalizing behavior problems: Group and individual differences. *Development and Psychopathology, 10*, 469-493.
- Fast Track Project. (1997). *Psychometric properties of the Social Health Profile*. Durham, NC: Duke University.
- Fraser, M. W. (1996). Aggressive behavior in childhood and early adolescence: An ecological-developmental perspective on youth violence. *Social Work, 41*, 347-361.
- Fraser, M. W., Abell, M. D., Galinsky, M. J., Schopler, J. H., & Hodges, V. G. (2001). *Strong families: A program of parent education*. Chapel Hill, NC: University of North Carolina, School of Social Work.
- Fraser, M. W., & Galinsky, M. J. (1997). Towards a resilience-based model of practice. In M. W. Fraser (Ed.), *Risk and resilience in childhood: An ecological perspective* (pp. 265-275). Washington, DC: NASW Press.
- Fraser, M. W., Nash, J. K., Galinsky, M. J., & Darwin, K. M. (2000). *Making choices: Social problem-solving skills for children*. Washington, DC: NASW Press.
- Fraser, M. W., Richman, J. M., & Galinsky, M. J. (1999). Risk, protection, and resilience: Toward a conceptual framework for social work practice. *Social Work Research, 23*, 131-143.
- Goldstein, H. (1995). *Multilevel statistical models* (2nd ed.). New York: Halsted.
- Hansen, D. J., Nangle, D. W., & Meyer, K. A. (1998). Enhancing the effectiveness of social skills interventions with adolescents. *Education and Treatment of Children, 21*, 489-513.
- Kelley, B., Huizinga, D., Thornberry, T., & Loeber, R. (1997). *Epidemiology of serious violence*. Washington, DC: U.S. Department of Justice, Office of Juvenile Justice and Delinquency Prevention.
- Koop, C. E., & Lundberg, G. D. (1992). Violence in America: A public health emergency. *Journal of the American Medical Association, 267*, 3075-3076.
- Littell, R. C., Milliken, G. A., Stroup, W. W., & Wolfinger, R. D. (1996). *SAS system for mixed models*. Cary, NC: SAS Institute.
- Loeber, R., Farrington, D. P., Stouthamer-Loeber, M., & Van Kammen, W. B. (1998). *Antisocial behavior and mental health problems: Explanatory factors in childhood and adolescence*. Mahwah, NJ: Lawrence Erlbaum.
- Nash, J. K. (1999). *Understanding and preventing youth violence: A pilot study of the Making Choices skills-training program*. Unpublished doctoral dissertation, University of North Carolina at Chapel Hill.
- National Institute of Mental Health. (1998). *Priorities for prevention research at NIMH*. Retrieved from <http://www.nimh.nih.gov/research/priorrpt/index.htm>
- Norusis, M. J. (2000). *SPSS 10.0: Guide to data analysis*. Upper Saddle River, NJ: Prentice Hall.
- Pettit, G. S. (1997). The developmental course of violence and aggression: Mechanisms of family and peer influence. *Anger, Aggression, and Violence, 20*, 283-299.
- Prothrow-Stith, D. B. (1995). The epidemic of youth violence in America: Using public health prevention strategies to prevent violence. *Journal of Health Care for the Poor and Underserved, 6*, 95-101.

- Rothman, J., & Thomas, E. J. (Eds.). (1994). *Intervention research: Design and development for human service*. New York: Haworth.
- Snyder, H. N. (2000). *Juvenile arrests 1999*. Washington, DC: U.S. Department of Justice, Office of Juvenile Justice and Delinquency Prevention. Retrieved from [http://www.ncjrs.org/html/ojjdp/jjbul2000\\_12\\_3/contents.html](http://www.ncjrs.org/html/ojjdp/jjbul2000_12_3/contents.html)
- Snyder, H. N., & Sickmund, M. (1999). *Juvenile offenders and victims: 1999 national report*. Washington, DC: Office of Juvenile Justice and Delinquency Prevention. Retrieved from <http://www.ncjrs.org/html/ojjdp/nationalreport99>
- U.S. Public Health Service. (2001). *Youth violence: A report of the U.S. Surgeon General*. Retrieved from <http://www.surgeongeneral.gov/library/youthviolence/sgsummary/summary.htm>
- Valente, E., & Dodge, K. A. (1997). Evaluation of prevention programs for children. In R. P. Weissberg, T. P. Gullota, R. L. Hampton, B. A. Ryan, & G. R. Adams (Eds.), *Establishing preventive services* (pp. 183-218). Thousand Oaks, CA: Sage.