Effect of a Structured Arts Program on the Psychosocial Functioning of Youth from

Low-Income Communities: Findings from a Canadian Longitudinal Study

Robin Wright, Lindsay John, and Stephen Ellenbogen
McGill University

David R. Offord and Eric K. Duku
McMaster University

William Rowe
University of South Florida

Abstract

This study reports on the longitudinal examination of a structured arts program for Canadian youth, aged 9 to 15 years, from low-income communities. Evaluated were the extent to which community-based organizations successfully recruited and retained youth in the program and whether they demonstrated improvement with respect to artistic ability (combination of theatre, visual and media arts) and psychosocial indicators. The results suggest successful recruitment and good retention rates. Multilevel growth curve analyses of observational and behavioral outcomes are presented. Observer ratings showed significant gains in artistic and social skills. Comparisons with matched controls using estimated linear propensity scores revealed a significant reduction in emotional problems for the intervention group. The overall conclusion is that youth from low-income communities benefit from structured arts programs.

Key words: Community-based arts programs, positive youth development, growth curve analysis, low-income communities, and emotional problems.

Effect of a Structured Arts Program on the Psychosocial Functioning of Youth from

Low-Income Communities: Findings from a Canadian Longitudinal Study

The importance of the arts on optimal youth development has been well argued

(Catalano, Berglund, Ryan, Lonczak, and Hawkins, 2002). However, the poorest Canadians have

limited exposure to these life-enriching experiences. Data from the National Longitudinal Survey

of Children and Youth (NLSCY) show that 65% of poor children have never played a musical

instrument, compared to 23% of middle-class children (Offord, Lipman, and Duku, 1998). Are

youths from low-income communities uninterested in artistic activities or are the arts programs

inaccessible to this segment of the population? Prior evaluations of arts programs targeted to

low-income or at-risk youth indicate that obtaining a sustained commitment can be challenging

(Clawson and Coolbaugh, 2001; Witt and Bradberry, 2000). In this study we report on the

National Arts and Youth Demonstration Project (NAYDP), which evaluated community-based

after-school arts programs (combination of theatre, visual and media arts) targeted to a low-

income population. Of particular interest were whether participants would demonstrate regular

and sustained attendance and whether the program would have an impact on their psychosocial

functioning.

Conceptual Orientation

From a program development perspective, the NAYDP was guided by the conceptual

framework of positive youth development. Lawrence (1998) operationalizes positive youth

development as an approach to working with youth, which assumes that they all engage in a

developmental process by which they seek to meet their needs and build competencies. The

development is considered positive since it departs from the deficit model by assuming that

youths have innate abilities and talents, and that they simply need opportunities to develop them.

This can best be achieved in environments that underscore strengths and nurture positive child/adult relationships (Lawrence, 1998). We agree with Peterson (2004) that the term positive youth development has been used indiscriminately to describe any program that involves youths. However, many of the tenets of positive youth development are reflected in the structure of the NAYDP. For example, the program aimed to engage youth in productive and life-enriching activities rather than correct or treat problems. This fact was underscored by both parents and program staff. However, it was considered important to monitor the behavioral indicators that are known to reduce quality of life. Art instructors and other support staff were screened for positive personal characteristics – a critical aspect of successful programs (Peterson, 2004). During staff training, the researchers stressed the importance of being intentionally inclusive, multicultural, and systematically non-discriminatory (Nicholson, Collins, and Holmer, 2004), and of fostering a supportive environment where youth would feel physically and emotionally safe. According to Catalano et al. (2002), positive youth development programs seek to achieve one or more of the following objectives: promoting social, emotional, cognitive, moral, and behavioral competencies; encouraging bonding; providing recognition for positive behavior and opportunities for prosocial involvement; and fostering resilience, self-determination, spirituality, self-efficacy, positive identity, belief in the future, and prosocial norms. With the exception of the prescriptions relating to moral competence, spirituality, and belief in the future, the NAYDP adhered to all these objectives.

<div align="center">Limitations of Existing Research</div>

There is an emerging literature on the positive effects of arts programs on youth. However, these programs are not homogeneous, and few have been rigorously evaluated. We will report on only those that have been subjected to some form of empirical evaluation. We can

divide these into three broad categories: school-based arts programs (Catterall, Chapleau, and Iwanaga, 1999; Upitis and Smithrim, 2003), after-school programs with the arts as one component among many (Jones and Offord, 1989), and interventions specifically designed to assess the effects of arts on high-risk youth (Clawson and Coolbaugh, 2001). Catterall, Chapleau, and Iwanaga (1999) analyzed American data from the National Educational Longitudinal Survey (NELS; 88) to ascertain the relationship between involvement in the arts and academic success. The main findings showed that students involved in school-based arts programs demonstrated increased creativity, lower dropout rates, increased social skills, and higher academic achievement. They also found that students from low socio-economic backgrounds who were involved in many artistic activities fared better on selected academic indicators than students from similar socio-economic backgrounds who were not involved in the arts. Winner and Cooper (2000), having reviewed studies making claims about the relationship between the arts and academic achievement, stated that much of this research is correlational in nature, and found that researchers sometimes go beyond their evidence to make claims about the association between arts and academic achievement. Upitis and Smithrim (2001) conducted a three-year Canadian study on the effects of the arts on academic achievement, in which they compared students in schools with arts curricula to control schools without arts curricula. Students in schools with arts curricula did significantly better on mathematical tests. However, once household income was included in a regression analysis, this finding was no longer significant.

After-school programs with artistic activities as one of the components have primarily targeted low-income and at-risk youth. In Canada, the effects of a sports, arts, and life-skills program (Jones and Offord, 1989) implemented in a large housing complex were rigorously evaluated with a comparable complex serving as a control. Using multivariate analyses, the study

found a clear and statistically significant reduction in criminal and delinquent behavior in the experimental housing complex. However, it showed no spillover effects on school performance and only minor, if any, effects on behavior in the home. In the YouthARTS Development Project (Clawson and Coolbaugh, 2001), the effects of arts programs targeting high-risk youths (e.g., with juvenile records) were examined up to 22 months after participation. Unfortunately, since numerous factors compromised the study, the authors could not unequivocally conclude that the programs were successful, although they did suggest that the findings were very encouraging.

The literature suggests that school-based arts programs have some effects on academic achievement and psychosocial functioning. On the other hand, there is little evidence of the efficacy of community-based arts program for youth living in low-income communities. The quality of evaluation of these programs makes it difficult to verify their success. A comprehensive review by the Rand Corporation of existing evaluations on the impact of arts-based programs for at-risk youth suggests that few provide good evaluations of their outcomes (Clawson and Coolbaugh, 2001).

To advance knowledge on the effects of community-based structured arts programs on youth from low-income communities, the NAYDP had three objectives: (a) evaluate the extent to which community-based organizations can successfully recruit, engage, and retain youths in a nine-month artistic endeavor centered around a theatrical performance or video production, (b) assess their in-program progress in terms of artistic and social skills development, and (c) determine whether involvement in arts programs is related to improved psychosocial functioning, thus attenuating behavioral and emotional problems such as conduct disorder, depression, hyperactivity, and low self-esteem.

Method

Site Selection and Participants

Data for this study are from the National Arts & Youth Demonstration Project, a quasi-experimental research design in which groups were matched rather than randomly assigned. This was a three-year longitudinal study to evaluate a community-based youth arts program implemented in five low-income communities across Canada. Site selection was based on the following criteria: a clear philosophy and mission, a well-defined infrastructure, financial stability, and credibility in the community. Sites had to agree to develop, implement, and monitor arts activities as prescribed in research protocol manuals. For example, they had to select arts activities that promoted the development of social skills and complied with data collection procedures. Reflecting an ethnically and culturally rich cross-section of the Canadian population, the selected sites were located in Montreal, Quebec; Toronto and a rural town in Ontario; Winnipeg, Manitoba; and Vancouver, British Columbia. For confidentiality reasons, sites will hereinafter be identified by a number from one to five. Local on-site staff from the same cultural community as the majority of participants were recruited for the project. Staff at each site included a lead arts instructor, two to three instructor assistants, and a research assistant. The lead arts instructor developed the curriculum under the guidance of the principal investigators. The research assistant coordinated the logistics of the arts program, ensured parental involvement by providing regular updates on youth absences and behavior, participated in program activities with youth, and collected data.

The principal investigators conducted information sessions at each site, meeting with potential artists, parents, representatives from local schools and police departments, and board members of arts organizations. This helped instill an atmosphere of community cooperation and

gave the investigators an opportunity to tap potential sources of assistance. After attending

orientation, data collection, and youth management training sessions, site personnel employed a

three-stage process to recruit parents, children, and youths to the program. The process

comprised: (a) a community mapping process to identify appropriate locations for community

outreach, (b) the development of active recruitment strategies such as setting up booths and/or

posting advertisements in local shopping malls, housing projects, schools, community-based

organizations serving parents or youth, ethno-specific organizations, and parks and recreation

centers, and (c) parent information sessions detailing the goals of the study.

Participants were recruited between July and August 2002. A total of 183 youths were

recruited. It took an estimated 30-45 hours to recruit a group for one site (30-35 youths). The

baseline data was collected in September 2002. Youth completed the self-report questionnaires

in their art classes. The research assistants met with the parents in their homes or other

convenient places to complete the self-report questionnaires. The program ran from the end of

September 2002 to the end of June 2003, and data was collected at the end of each of the three

terms: December 2002, March 2003, and June 2003. A six-month follow-up assessment was

carried out in December 2003. To facilitate the children's ongoing contact with the arts

organizations and their peers, and to maintain program effects over time, two booster sessions

were conducted in the fall of 2003.

The Arts Program

The arts programs were free. They were offered as twice-weekly ninety-minute sessions

for a total of 37 weeks. Transportation to and from the program was provided free of charge, as

were snacks. The program was divided into three increasingly challenging and cumulative terms.

Though no restrictions were placed on the arts medium to be taught, all sites followed a similar

structure. The first term of the program generally consisted of an exploration of several arts media: improvisation, painting, mask making, music, scriptwriting, filming, etc. The focus was on self-exploration, having fun, and developing positive group dynamics (e.g., participants were urged to respect one another and resolve conflicts non-violently). Following this initial exposure to many facets of the arts, the second term of the program allowed youth to pursue more challenging goals in their chosen medium. This required greater teamwork. The last term pulled together all the youths' acquired artistic and teamwork skills to stage a theatrical performance or showcase a video production in their community.

Measures

In order to use youth from the National Longitudinal Survey of Children and Youth (NLSCY) as a comparison/control group for estimating the NAYDP program effects, we incorporated identical measures in our study. The NLSCY is an ongoing longitudinal household survey designed to monitor the development and well-being of a representative sample of Canadian children. Data on children aged up to 11 years was first collected in 1994, and every two years thereafter until the sample reached the age of 25. The progress of the arts program participants was evaluated using four observational measures. These were originally used in the NLSCY to gather teacher ratings of student classroom behavior. Based on observation notes taken during every class, research assistants rated participants twice per term for a total of six assessments. The following measures were used: *Joyful participation in activities* – three items measuring enjoyment of and engagement in activities. A sample item is, "This child/youth fully participates in activities." *Social skills development* – 12 items measuring self-control, communication, respect, cooperation, and problem solving. A sample item is, "This child/youth demonstrates cooperative play with other youth." *Art skills development* – two items measuring

whether youth meets goals and shows improvement. A sample item is, "This child/youth shows improvement in his/her art skills." *Task completion* – seven items measuring listening skills and work habits. A sample item is, "Completes work on time." All four scales have a five-point response category (never, rarely, sometimes, often, always). We assessed the internal consistencies of the scales using Cronbach's Alpha (α). If items within a scale purport to measure a common entity, they must be internally consistent. Hence, they should be correlated with one another. Nunnaly (1978) has indicated 0.7 as an acceptable reliability coefficient. The alphas for the observational scales were consistently high across times of measurement, ranging from .89 to .92 for *Joyful participation in activities*, .95 to .97 for *Social skills development*, .90 to .96 for *Art skills development*, and .95 to .97 for *Task completion.*

We used five self-report behavioral outcome measures, namely, *Conduct* problems, a seven-item scale measuring bullying, getting into fights, and vandalism. A sample item is, "I get into many fights." *Emotional problems*, an eight-item scale measuring unhappiness, depression and anxiety. A sample item is, "I am not as happy as other people my age." *Self-esteem* is a five-item scale. A sample item is, "In general, I like the way I am." *Prosocial behavior* is a ten-item scale measuring the children's empathic and helpful behavior towards other children. A sample item is, "I will try to help someone who has been hurt." *Hyperactivity/attention deficit* is an eight-item scale measuring ability to concentrate, restlessness, and fidgeting. A sample item is, "I can't concentrate, and I can't pay attention for long." All were measured on a three-point response category (never or not true, sometimes or somewhat true, often or very true) except for self-esteem, which had a five-point response category (false, mostly false, sometimes false/sometimes true, mostly true, true). Except for *Hyperactivity/attention deficit,* the alphas were consistently high across times of measurement, ranging from .86 to .91 for *Conduct*

problems, .76 to .89 for *Emotional problems*, .85 to .92 for *self-esteem*, and .83 to .92 for

Prosocial behavior. The internal consistency for *Hyperactivity/attention deficit* ranged from 0.10

to 0.87. The hyperactivity measure was not included in the analyses due to an unacceptably low

Cronbach's alpha.

The hyperactivity, conduct problem, and emotional problem measures were taken from

the Ontario Child Health Study (Boyle et al., 1987). Although all measures were taken from the

NLSCY survey, the self-esteem measure was first introduced in the Marsh Self-Description

Questionnaire (Statistics Canada, 1995), and the prosocial scale comes from the Montreal

Longitudinal-Experimental Study (Tremblay et al., 1992). We rescaled all observational and

outcome measures to a range of 0 to 10. The person most knowledgeable (PMK), typically the

mother, completed questionnaires to gather socio-demographic information and assess *family*

functioning (α = .83), a twelve-item scale measuring problem-solving, communication, roles,

affective involvement, and behavioral control. This measure is a subscale of the McMaster

Family Assessment Device (Byles, Byrne, Boyle, and Offord, 1988). We used baseline data for

Family functioning only as a covariate matching estimator

Control Group

We selected the control group using propensity score matching. The propensity score is

defined by Rosenbaum and Rubin (1983) as the conditional probability of assignment to a

particular treatment given a vector of observed covariates. An important feature of matching on

the estimated linear propensity scores is that it does not require close pairwise matches

(Rosenbaum and Rubin, 1983). To perform the propensity score matching, we combined the

NAYDP sample with a selected sample of NLSCY youths aged 9 to 15 years who had at least

three measurement points. To obtain more precise matches, we calculated age in months. We

then performed a logistic regression analysis, with the outcome variable defined as 1 =

intervention group (NAYDP) and 0 = control group (NLSCY). The covariates entered as

predictors in the logistic model were baseline scores of child-rated conduct problems, emotional

problems, hyperactivity/attention deficit, self-esteem and prosocial behavior, as well as child's

gender, PMK education, PMK marital status, household income, and family functioning. The

logistic regression model collapses the specified covariates into a single variable, namely the

predicted probability of being in the intervention group, also referred to as propensity score. The

predicted probabilities obtained from the logistic model were saved as propensity scores (one for

each child). We then matched each NAYDP participant with the NLSCY respondent with the

closest propensity score at baseline. We compared the mean scores of the NAYDP and NLSCY

subjects for all covariates using independent t-tests and odds ratios. The two cohorts were very

similar on all covariates: odd ratios ranged from 0.97 to 1.04 and significance levels for the t-

tests exceeded 0.25.

Analytical Strategy

We performed statistical analyses for demographic information and assessing internal

consistencies of the scales using the SPSS statistical package. We addressed the amount and

direction of changes in the observational and behavioral outcomes with a multi-level growth

curve analysis using a hierarchical linear modeling software package (HLM 5.0; Raudenbush,

Bryk, Cheong, and Congdon, 2001). Multi-level growth curve analysis is a method of assessing

the relationships among variables that violate the assumptions of data independence (Willett,

Singer, and Martin, 1998). Since the data from the NAYDP are not only repeated measurements

on the same subject but also nested data, i.e., youths nested within sites within neighborhoods,

they are correlated. This analytic strategy allowed us to include partial response data. We took

into consideration that not all the children, e.g. dropouts and new recruits, were assessed at the same time or for the same number of occasions. Growth curve models make use of all the available data for a given subject, even if data for that subject is not available for all occasions. Growth curve models also impose no restrictions on measurement schedules.

Results

Sample Description

The sample consisted of 64.5% or 118 girls (F) and 35.5% or 65 boys (M). When they started the program, 27% (36 F, 13 M) of the youth were 9-10 years old, 60% (71 F, 39 M) were 11-12 years old, 10% (10 F, 9 M) were 13-14 years old, and 3% (2 F, 3 M) were 15 years old. The sample was 59% White, 26% Aboriginal, 11% Black, 6% Asian, 5% Latino, and 4% other. The family structure was 38% married, 14% single, 24% separated or divorced, 20% living in common-law relationships, and 4% widowed. Of the reporting parents, 40% were high school dropouts, 32% had received welfare in the last 12 months, and 40% reported earning a total household income of less than $20,000/year.

Comparisons with national statistics (National Council of Welfare, 2004; Statistics Canada, 2001, 2001a, 2002, 2004) indicate that the NAYDP was successful in recruiting youth from low-income communities. Although different measuring techniques were employed, a few reasonable points of comparison were possible (NAYDP average in brackets): 75% (58%) of Canadian families were formed with married or common-law partners in 2001; 7% (32%) of those under 65 years of age were welfare recipients in 2002; 20% (40%) of families earned less than $20,000/year in 2002; and 33% (40%) of those 16 and over did not complete high school in 2001.

Sample Loss and Attendance

Following recommendations by Chaput, Little, and Weiss (2004), we expanded the concept of attendance to include duration (number of terms completed) and intensity of participation (number of sessions attended per term). Table 1 presents the attendance statistics. Of the 159 people who started at the beginning of the nine-month program, 58 or 36% dropped out of the program. The attrition rate was 37% for the 183 youth who participated at some point in the program, including late joiners. With respect to duration of attendance, 84% of participants completed at least one term, and 72% of those capable of completing at least two terms (i.e., excluding Term 3 joiners) did so. The two reasons most often cited for dropping out were loss of interest in the program (26%) and moving from the area (18%). Attendance rate was over 80% for all terms. Overall, this indicates a stable and committed engagement in the program by the majority of participants.

[Insert Table 1]

The 68 participants who dropped out of the program were compared with the 115 participants who completed Term 3. Using chi-square analyses and independent sample t-tests, we found no significant differences with respect to the participants' gender or age, nor with respect to their parents' marital status or being on welfare. Youth were more likely to drop out if their PMK did not graduate from high school, $\chi2$ (1, N= 172) =4.31, p=.04 or if their families earned less than \$20,000, $\chi2$ (1, N= 174) =5.63, p=.02. It should be noted that 57% of the participants whose PMK did not graduate from high school completed the program, as did 56% of the children in families earning less than \$20,000. Levels of youth-reported self-esteem, depression, and prosocial behavior were not significantly different between dropouts and completers. Dropouts tended to have higher scores for conduct problems ($M = 2.28$, $SD = 2.45$) compared to those who completed the program ($M = 1.96$, $SD = 2.36$), t $(179) = 2.75$ (two-

tailed), $p = .07$, $d = .65$. We also found site differences: $\chi2$ (4, N = 183) = 13.18, p = .01. Site four

had the highest dropout rate (57%), followed by site two (43%), three (33%), one (27%), and

five (21%).

Observational Outcomes

The analysis for evaluating the youths' in-program behavior was achieved by regressing

each repeatedly measured outcome on time itself. Time was represented by a dummy variable for

the six measurement points. To measure change over time with repeated measurement data, a

pair of linked statistical models must be specified (Willett, 1988). Hence, we employed a two-

level hierarchical model for observational outcomes using the following equations:

Level-1 model: $Y = B_O + B_1 \text{(time)} + r$

Level-2 model: $B_O = G_{OO} + G_{O1}* \text{(Female=1)} + G_{O2}*\text{(Age centered)} + U_O$

$\quad\quad\quad\quad B_1 = G_{10} + G_{11}* \text{(Female=1)} + G_{12}*\text{(Age centered)} + U_1$

where Y is the outcome variable and B_O and B_1 are the coefficients for the intercept (baseline

scores) and growth trajectory (slope), respectively. In hierarchical linear modeling, the intercept

(B_O) and growth (B_1) coefficients from the Level-1 model become the outcome variables for the

Level-2 model. G_{OO} is the Level-2 intercept for the average baseline scores, and G_{O1} and G_{O2}

assess the extent to which average baseline scores of the Level-1 model vary as a function of

gender and age. G_{10} is the Level-2 intercept for average growth trajectories (slope) of the level-1

model. G_{11} and G_{12} assess the extent to which average slope of the Level-1 model also varies as a

function of gender and age. The residual or error terms U_O, U_1 are measures of the variation

around the estimates of the intercept and slope, respectively, and r is the unexplained variance in

the model. We chose gender as a predictor for the Level-2 model, since there is some discussion

as to whether males demonstrate equal commitment to arts programs (Upitis and Smithrim,

2001). We chose age as a predictor because policy-makers and funding agencies raised concerns about the optimal age for participation in these programs.

For ease of interpretation, we centered age at the grand mean (11.3 years). Centering around the grand mean expresses the individual scores as deviations from the grand mean. Hence, the estimated coefficients indicate the amount of change in specific outcomes for each unit of change in age from the grand mean. The results are shown in Table 2. Estimated modeled parameters were divided into fixed and random effects. The fixed effects give the sample averages for the intercept and growth trajectories for the two-level model. The intercept and trajectory of the fitted Level-1 and Level-2 models were allowed to vary randomly. Referred to as random effects, they show the deviations from the average intercepts and average growths of the Level-1 and Level-2 models. In addition, the beta coefficients and their standard errors are presented under fixed effects. Level of significance, represented by a t-ratio, is the coefficient divided by its standard error.

An examination of the fitted parameters (intercept and growth trajectory) for Level-1 revealed that the average intercepts of the four observational outcomes were all above scale midpoints and all significantly different from zero. The average participant growth trajectories for the observational outcomes were all significantly different from zero. Estimated parameters for the added predictors of gender and age are presented under the Level-2 fixed effects. Controlling for age, females started at a higher score than males on all four observational outcomes. Controlling for gender, average intercept for age (centered at 11.3 years) was statistically significant only for social skills development. The older the youth, the more social skills manifested in the arts programs. This is not unexpected, since older children are usually more socially adept than younger children. The estimated Level-2 parameters showed that the

linear trajectories for the observational outcomes did not differ by gender or age. This tells us that, no matter what the average starting points on the four observational outcomes, youths of different ages and gender progressed at similar rates.

The random effects showed significant variations in average intercept and growth trajectories on all observational outcomes for both the within- and between-subject models. Unexplained variance for the Level-1 and Level-2 models was similar across all observational outcomes. This is to be expected, since the two coefficients of gender and age were not statistically significant to explain the parameter variance within the Level-1 model.

<div align="center">**[Insert Table 2]**</div>

Behavioral Outcomes

Unlike the modeling for observational outcomes, the modeling to compare the NAYDP and NLSCY participants regresses the repeated behavioral measurement on age (in months). Age instead of time was used as the independent variable, because the NAYDP and NLSCY children were not assessed for the same number of occasions or at similar intervals. Hence, the growth trajectory is based on the child's age. For example, a child who was nine years old (108 months) at entry into the NAYDP program had five measurement points (at 108, 111, 114, 117, and 123 months). An NLSCY respondent entering the survey at age nine had four repeated measurements on the covariates at the end of cycle 4 of the survey (108, 132, 156, and 180 months). The question of interest is how the growth trajectories of the NAYDP participants compared to those of the NLSCY controls on the four selected outcomes. The model for the behavioral outcomes is represented by the following equations:

Level-1 model: $Y = B_O + B_1 * (\text{Age centered}) + r$

Level-2 model: $B_O = G_{OO} + G_{O1} * (\text{NAYDP}=1) + U_O$

$$B_1 = G_{10} + G_{11}*(NAYDP=1) + U_1$$

Estimated average intercept and average growth trajectories for the intervention

(NAYDP) and control (NLSCY) groups are shown in Table 3. Since an intercept can be defined

as the average score for a subject who is coded zero on the covariate, G_{00} and G_{10} are the

intercept and growth trajectory for the NLSCY participants. G_{01} measures the mean difference

between NAYDP and NLSCY subjects. The program effect on the rate of change or growth

trajectories for the two groups is estimated by G_{11}. For the behavioral outcomes, age was

centered at 12. The following outcomes are reported: conduct problems, emotional problems,

self-esteem, and prosocial behavior. Since we were interested in comparing only the growth

trajectories of the two groups, the average intercept (B_0) and growth (B_1) for the entire sample

are not reported.

As can be seen from Table 3, with the exception of conduct problems, average intercept

scores for all outcomes were lower for the NAYDP participants. Growth trajectories for the

outcomes of conduct disorder, self-esteem, and prosocial behavior, although significantly

different from zero, were similar for both groups. The only significant difference between the

two groups was for emotional problems. The NAYDP participants showed a significant decrease

in emotional problems, while the NLSCY participants showed a non-significant decrease.

In terms of random effects, there were significant variations around the intercept for all

four behavioral outcomes, although they were smaller than those reported for the observational

outcomes. The only non-significant variation around the growth trajectories was for conduct

problems. The variance around the growth trajectories, although significant, was not very

substantial, ranging from 0.00004 to 0.0005.

[Insert Table 3]

Discussion

The first objective of the study was to evaluate the extent to which community-based organizations can successfully recruit, engage and retain youth from low-income communities in structured artistic activities. The high attendance rates suggest that youth from low-income communities can commit to an intensive arts program once obstacles such as cost and lack of transportation are removed. Moreover, the NAYDP youth have shown a significant improvement from baseline to final assessment in the "joyful participation" observational measure. The participation rates are very encouraging, considering that one of the characteristics of low-income families and communities is a high rate of transience – a factor that accounted for 17% of the NAYDP dropouts.

A comparison of program dropouts and completers reveals that those who persisted in the program were more likely to come from higher-income families in which the PMK had a high school diploma. Although it could be argued from these findings that a sustained engagement was obtained solely from participants without socio-economic or other hindrances, a more thorough examination of the results suggests that this was not the case. First, we must keep in mind that the majority of youth in all demographic categories completed the program. Furthermore, PMK marital status did not significantly affect attrition, nor did being on welfare. That children of single parents were equally engaged in the program is important, because this group has often been identified as particularly in need of quality after-school activities (Gottfredson, Gerstenblith, Soulé, Wormer, and Lu, 2004).

The second objective of the study concerned the youth's in-program progress with respect to artistic and social skills development. The significant improvements on all observational measures from baseline to final assessment indicate that structured arts programs provide an

opportunity for youth to develop important communication, cooperation, conflict resolution, and teamwork skills. These findings corroborate the assertion that every child can be brought to a certain level of arts literacy (Pitman, 1998). Proponents of positive youth development (Catalano et al., 2002; Lawrence, 1998) believe that this can only be achieved in environments that underscore strengths and nurture positive child/adult relationships. Females start higher than males on most of the observational measures. However, both improved at the same rate. It is speculated that the supportive and non-competitive atmosphere ensured that children with lesser artistic abilities and social competencies—boys in this case—would not feel discouraged. This finding is promising, because Upitis and Smithrim (2003) had found that females tended to enjoy arts activities more than boys did.

The third objective of the study was to determine whether the arts programs had an impact on the participants' psychosocial functioning. Compared to the control group, the results indicated a statistically significant decrease in emotional problems. This finding is congruent with Richardson, Radziszewska, Dent, and Flay (1993), who found a relationship between a lack of after-school supervision and depressed mood. By introducing a new social environment, the NAYDP may have contributed to an increase in the quantity and quality of the participants' social network. It could be that the program stimulated the core elements for personal growth mentioned in the introduction (Catalano et al., 2002). The program may have fulfilled a need for greater stimulation, which in turn boosted their spirits. Finally, one could argue that the contact with caring and non-judging adults provided the youth with new opportunities to share their emotional difficulties.

The growth trajectories of the NAYDP and NLSCY participants for conduct problems, prosocial behavior, and self-esteem were similar. It should be noted that children with behavioral

difficulties were not specifically targeted for this study, which explains why the mean intercept seems relatively low (1.59 out of 10). The 2:1 ratio of females to males must also be considered, given that problems with conduct and hyperactivity are more prevalent among boys (Willms, 2002). With these in mind, the present study cannot address the broader question of whether arts instruction can prevent conduct problems, since the participants were simply not "at-risk" enough. The decrease in prosocial behavior for both the NAYDP and NLSCY participants was unexpected. The speculation is that some of the items on that scale, such as "helps other children who are sick" or "comforts a child who is crying," might be considered less socially desirable at certain stages of adolescence. The term "child" may also be problematic; substituting the term "classmate" may be advisable in future studies.

Considering the improvement in emotional well-being, it seems somewhat odd that there was no concurrent increase in self-esteem. If the participants felt less depressed, unhappy and anxious, should that not be a reason to feel better about themselves? Fleming and Offord (1990) reported an association between self-esteem and depressive disorder. It could be that the mechanisms that regulate self-esteem (the image we have of ourselves) are more ingrained than those that regulate emotional well-being (how we feel at the moment).

Study Limitations and Recommendations for Future Research

We used a matched sample design to ensure that the groups of the present study were very similar with respect to socio-economic status and psycho-social functioning. However, there may be relevant characteristics for which matching strategies are unable to control. The most obvious difference is that the intervention group agreed to take part in the program, whereas the comparison group may or may not have chosen to do so. As a result, the intervention group may have been more primed for personal growth. However, most of the members of the

intervention group were urged to take part in the program, and it seems rather unlikely that they would have joined a performing arts program had recruiters not actively solicited their participation. While a design involving a randomized assignment of youth to intervention and control groups remains the strongest test of effectiveness, a matched design is nevertheless far superior to having no control at all. Another limitation involves the validity of the observational measures. Since the research assistants who performed the assessments may have been aware of the desired outcomes, rater bias is a concern.

Even though the NAYDP participants demonstrated little evidence of conduct problems, the research assistants' observational assessments and interviews with the youths, parents, and project staff suggest that their conduct did improve. In addition, to better understand the effects of community arts programs on youth, future research should focus on: (a) a more targeted population of at-risk youth, (b) comparing different arts media, and (c) comparing different levels of duration and intensity of arts programming.

References

Boyle, M. H., Offord, D. R., Hofmann, H. G., Catlin, G. P., Byles, J. A., Cadman, D. T.,

Crawford, J. W., Links, P. S., Rae-Grant, N. I., and Szatmari, P. (1987). Ontario Child

Health Study. I. Methodology. *Archives of General Psychiatry, 44*(9), 826-831.

Byles, J., Byrne, C., Boyle, M. H., and Offord, D. R. (1988). Ontario Child Health Study:

reliability and validity of the general functioning subscale of the McMaster Family

Assessment Device. *Family Process, 27*(1), 97-104.

Catalano, R. F., Berglund, M. L., Ryan, J. A. M., Lonczak, H. S., and Hawkins, J. D. (2002).

Positive youth development in the United States: Research findings on evaluations of

positive youth development programs. *Prevention & Treatment, 5*(15), 1-112.

Catterall, J. E., Chapleau, R., and Iwanaga, J. (1999). *Involvement in the arts and human

development: general involvement in music and theatre arts*. Los Angeles: Imagination

Project, UCLA School of Education and Information Studies.

Chaput, S. S., Little, P., and Weiss, H. (2004). *Understanding and Measuring Attendance in

Out-of-School Time Programs*. Cambridge, MA: Harvard Family Research Project.

Retrieved November 2, 2004, from http://www.gse.harvard.edu/hfrp/projects/

afterschool/resources/ issuebrief7.html

Clawson, H. J., and Coolbaugh, K. (2001). *The YouthARTS Development Project. Juvenile

Justice Bulletin*. Washington: The Office of Juvenile Justice and Delinquency. Retrieved

November 29, 2004, from http://www.ncjrs.org/html/ojjdp/2001_5_2/contents.html.

Fleming, J. E., and Offord, D. R. (1990). Epidemiology of childhood depressive disorders: a

critical review. *Journal of American Academic Child and Adolescent Psychiatry, 29*(4),

571-580.

Gottfredson, D. C., Gerstenblith, S. A., Soulé, D. A., Wormer, S. C., and Lu, S. (2004). Do
after school programs reduce delinquency? *Prevention Science, 5*(4), 253-266.

Jones, M. B., and Offord, D. R. (1989). Reduction of antisocial behaviour in poor children by
nonschool skill-development. *Journal of Child Psychology & Psychiatry, 30*(5), 737-750.

Lawrence, D.H. (1998). Positive youth development promotes pathways of growth. *Common
Ground* 15: 9.

National Council of Welfare. (2004). *Fact sheet: welfare recipients*. Retrieved November 9,
2004, from http://www.ncwcnbes.net/htmdocument/principales/numberwelfare_e.htm

Nicholson, H.J., Collins, C. and Holmer, H. (2004) Youth as people: the protective aspects of
youth development in after school settings. *ANNALS, AAPSS*, 591, 55-71.

Nunnaly, J. (1978). *Psychometric theory*. New York: McGraw-Hill.

Offord, D. R., Lipman, E. L., and Duku, E. K. (1998). Sports, the arts and community
programs: Rates and correlates of participation. (Working Paper W-98-18E). Hull, QC:
Applied Research Branch, Strategic Policy, Human Resources and Development Canada

Peterson, C. (2004, January). Preface. Positive Youth Development: Realizing the Potential of
Youth. *ANNALS, AAPSS*, 591, 6-12.

Pitman, W. (1998). *Learning the arts in an age of uncertainty*. Toronto: Arts Education
Council of Ontario.

Raudenbush, S., Bryk, A., Cheong, Y.K and Congdon, R. (2001). Hierarchical Linear and
Nonlinear Modeling, Rel. HLM 5.05. Lincolnwood, IL: Scientific Software International.

Richardson, J. L., Radziszewska, B., Dent, C. W., and Flay, B. R. (1993). Relationship
between after-school care of adolescents and substance use, risk taking, depressed mood,
and academic achievement. *Pediatrics, 92*(1), 32-38.

Rosenbaum, P. R., and Rubin, D. B. (1983). The central role of the propensity score in observational studies for casual effects. *Biometrika, 70*(1), 41-55.

Statistics Canada. (1995). *National Longitudinal Survey of Children: overview of survey instruments for 1994/95, data collection, cycle 1*. Retrieved November 11, 2004, from http://www.statcan.ca/english/freepub/89F0078XIE/96001.pdf

Statistics Canada. (2001). *Census families in private households by family structure and presence of children, by provinces and territories*. Retrieved November 9, 2004, from http://www.statcan.ca/english/Pgdb/famil54a.htm

Statistics Canada. (2001a). *Population 15 years and over by highest degree, certificate or diploma, by provinces and territories*. Retrieved November 9, 2004, from http://www.statcan.ca/english/Pgdb/educ41a.htm

Statistics Canada. (2002). *Analysis of income in Canada*. Retrieved November 9, 2004, from http://www.statcan.ca/english/freepub/75-203-XIE/00002/part7.htm

Statistics Canada. (2004). *Population by sex and age group*. Retrieved November 9, 2004, from http://www.statcan.ca/english/Pgdb/demo10a.htm

Tremblay, R. E., Vitaro, F., Bertrand, L., LeBlanc, M., Beauchesne, H., Boileau, H., and David, L. (1992). Parent and child training to prevent early onset of delinquency: The Montreal Longitudinal-Experimental Study. In J. McCord and R. E. Tremblay (Eds.), *Preventing antisocial behaviour: interventions from birth through adolescence* (pp. 117-138). New York: Guilford Press.

Upitis, R., and Smithrim, K. (2003). *Learning Through the Arts: National Assessment 1999-2002*. Toronto: The Royal Conservatory of Music.

Willett, J. B. (1988). Question and answers in the measurement of change. In E. Rothkopf

    (Ed.), *Review of Research in Education (1988-89)* (pp. 345-422). Washington, DC:

    American Education Research Association.

Willett, J.B., Singer, J.D., and Martin, N.C. (1998). The design and analysis of longitudinal

    studies of development and psychopathology in context: Statistical models and

    methodological recommendations. *Development and Psychopathology,* 10, 395-426 .

Willms, D. J. (2002). The prevalence of vulnerable children. In D. J. Willms (Ed.), *Vulnerable*

    Children (pp. 45-69). Edmonton: The University of Alberta Press.

Winner, E. and Cooper, M. (2000). Mute those claims: No evidence (yet) for a causal link

    between arts study and academic achievement. *The Journal of Aesthetic Education*, 34(3-

    4), 11-75.

Witt, P. A., and Bradberry, E. K. (2000). *Evaluation of the totally cool, totally art program.*

    Texas: Austin Parks and Recreation Department. Retrieved November 8, 2004, from

    http://www.rpts.tamu.edu/Faculty/Witt/conpubs/TCTA00.PDF

Author Note

Dr. David R. Offord passed away in April 2004 prior to the completion of the present paper.

Table 1

Participation by Term

	Term 1	Term 2	Term 3	Total
Joiners	159	18	6	183
Dropouts	26	25	17	68
Average attendance*	84%	80%	83%	82%
Percent attending over 50% of participation days*	93%	91%	95%	93%
Percent attending over 75% of participation days *	79%	71%	77%	76%

Note: Term = 24 participation days

*excluding participation days missed prior to joining and after dropping out

Table 2

Growth Curve Modeling of Observational Outcomes of In-program Youth Behavior

	Outcomes							
	Participation		Artistic skills development		Social skills development		Task completion	
Fixed Effects	*Beta*	*SE*	*Beta*	*SE*	*Beta*	*SE*	*Beta*	*SE*
Level-1 model								
Avg. intercept B_O	6.56^{**}	.16	6.30^{**}	.18	6.85^{**}	.16	6.81^{**}	.18
Avg. growth B_1	$.19^{**}$	.04	$.30^{**}$	.04	$.20^{**}$	.03	$.22^{**}$	.04
Level-2 model								
Avg. intercept G_{OO}	6.11^{**}	.29	5.70^{**}	.33	6.08^{**}	.30	6.09^{**}	.37
Female G_{O1}	$.71^{*}$	.34	$.92^{*}$	.39	1.12^{*}	.35	1.06^{*}	.43
Age (centered) G_{O2}	-.04	.12	.18	.13	$.32^{*}$	.12	.17	.13
Avg. growth G_{1O}	$.22^{*}$	.07	$.35^{**}$	.08	$.25^{*}$	.07	$.29^{**}$	
Female G_{11}	-.05	.09	-.08	.10	-.07	.08	-.11	.09
Age (centered) G_{12}	-.02	.03	-.007	.03	.004	.03	.03	.03
Random Effects	*Variance*		*Variance*		*Variance*		*Variance*	
Level-1 model								
Avg. intercept U_O	3.13^{**}		4.05^{**}		3.62^{**}		4.50^{**}	
Avg. growth U_1	$.11^{**}$		$.14^{**}$		$.11^{**}$		$.14^{**}$	
Level-1 effect, r	1.39		1.80		.75		1.08	

Table 2 (continued)

Random Effects	Variance	Variance	Variance	Variance
Level-2 model				
Avg. intercept *UO*	3.03**	3.85**	3.26**	4.27**
Avg. growth *U*1	.11**	.15**	.11**	.14**
Level-2 effect, *r*	1.39	1.80	.75	1.09

Note. Scale ranges from 0 (least improvement) to 10 (most improvement).

* *p* < .05. ** *p* < .01.

Table 3

Growth Curve Modeling of NAYDP Participants and NLSCY Matched Controls

	Outcomes							
	Conduct problems		Emotional problems		Self-esteem		Prosocial behavior	
Fixed Effects	Beta	*SE*	Beta	*SE*	Beta	*SE*	Beta	*SE*
Avg. intercept								
NLSCY G_{OO}	1.22^{**}	.10	2.52^{**}	.13	7.77^{**}	.11	6.73^{**}	.117
NAYDP G_{O1}	1.59^{*}	.15	2.47	.18	7.68	.17	6.43	.185
Avg. growth								
NLSCY G_{10}	$-.01^{*}$	.002	-.006	.005	$-.02^{**}$	.004	$-.018^{**}$	.003
NAYDP G_{11}	$-.01^{*}$	.002	$-.023^{*}$	.008	$-.02^{**}$	.004	$-.018^{**}$	.003
Random Effects	*Variance*		*Variance*		*Variance*		*Variance*	
Intercept U_O	1.47^{**}		2.00^{**}		1.67^{**}		2.22^{**}	
Growth U_1	.0003		$.0005^{*}$		$.0002^{*}$		$.00004^{*}$	
Level-1 r	1.78		2.10		2.30		2.37	

Note. Scale ranges from 0 (least problems) to 10 (most problems) for emotional and conduct problems, and from 0 (lowest score) to 10 (highest score) for self-esteem and prosocial behavior.

$* p < .05. ** p < .01.$