

The full report of the HIPP evaluation is available from the program evaluator: Copper Coggins CCogg@bellsouth.net or from the National Coordinator for HIPP: Kathryn Liss kliss@afsc.org.

HELP INCREASE THE PEACE MIDDLE SCHOOL PROGRAM EVALUATION

Summary Report

Purpose

This field experiment evaluates the impact of the Help Increase the Peace Program on aggressive behavior and victimization of sixth grade African American students in a middle school located in a high poverty, high crime urban neighborhood. The purpose of this study as a summative evaluation is to use the strongest possible research design to assess program effectiveness and thus aid adopting and funding agencies to determine the potential of the program in this context.

Program

The Help Increase the Peace Program (HIPP or HIP Program) is a project of the American Friends Service Committee that is centered on a 12 to 18 hour conflict resolution training intervention. For complete program information contact the American Friends Service Committee POB 73008 Washington, DC 20056 or see the AFSC website at <http://www.afsc.org/resources/items/hipp-manual.htm>. HIPP workshops were originally designed for middle and high school students in response to the problem of youth violence. The program has been disseminated widely and been positively received by participants. However, the program needed a strong comparative evaluation to test its effectiveness as a violence prevention intervention.

In the summer of 2004 the HIPP program coordinator reached an agreement with a public middle school to be a pilot site for an experimental evaluation and with the program evaluator to conduct an independent evaluation of the HIPP program at this site as part of her doctoral dissertation in Psychology in the Public Interest. The research was approved by the Institutional review board of the North Carolina State University and by the Baltimore City School District. The research was funded by a grant from the Maryland Mediation and Conflict Resolution Office (MACRO). The intervention took place as a series of 8 weekly two hour workshops for 6th grade students at the research site in the fall of 2004.

Goals for HIPP identified by the program developer and the pilot evaluation site are the following:

1. To decrease incidence of aggression and victimization among students.
2. To improve students' attitudes toward interpersonal conflict (reducing students' attraction to escalation and retaliation while increasing their attraction to peaceful problem solving options to resolve disputes).
3. To increase students' self-confidence in being able to "keep the peace."
4. To increase students' conflict resolution knowledge.

Method

This outcome evaluation is a true experiment: Students who volunteered for the HIP Program and whose parents did not object to their participation were randomly assigned (stratified by sex) either to be in the HIP Program or to be in a control group. The setting is an urban middle school with 99% African American students located in a high poverty, high crime neighborhood of an eastern U.S. city. Only African American 6th grade students were in the sample. Both treatment and control groups completed a pretest questionnaire one week before the HIPP workshops started and one week after the intervention ended. Self-report measures of behavior, attitude and self-efficacy as well as a curriculum knowledge test were included on the questionnaire to assess achievement of the four program goals listed above.

Only those behavior and attitude scales determined to have an acceptable reliability (above 0.60 Cronbach alpha) and confirmed (by factor analysis) as valid measures of the intended concept with this sample were used in analysis. Boys and girls responded so differently to some of the self-report measures that it would have not been appropriate to use the same measures for both sexes. The following outcomes could be measured with acceptably valid and reliable measures: physical aggression (for both boys and girls), victimization by peers (for boys only), self-efficacy for peaceful conflict resolution (for both boys and girls) attitudes supporting non-violence (for girls only) attitudes supporting aggression (for boys only). Knowledge of HIPP curriculum concepts was also measured for boys and girls using a 10 item multiple choice test.

The impact of the HIPP program was assessed by multiple regression analyses that statistically controlled for pretest scores and sex where relevant. The study also investigated the possibility that the program is received differently by girls and boys (Sex x Treatment interaction).

An underlying assumption of conflict resolution training (and other school prevention curricula) is that by giving students information and skills and by changing their attitudes they may be induced to make positive changes in behavior. The

behavioral outcomes (in this case victimization and aggressive behavior) are said to be mediated by the process variables (in this case self-efficacy, attitudes and knowledge). The mediation of behavioral outcomes by the process variables can be assessed statistically and was addressed in this study where warranted.

Student volunteers completing a useable pretest (41 girls, 32 boys) were randomly assigned to condition stratified by sex. Attrition was low (3/73 or 4%). Of the 73 participating students, one (a girl in the HIP Program) transferred out of district and was not available for posttesting. Data from the 72 students for whom there were matched pre- and posttests were used in refining measurement scales. According to the trainer's attendance records, two students (1 girl, 1 boy) who were assigned to the control group were inadvertently allowed to receive the full 8 workshop dose of the HIP Program. These students were removed from the sample used for hypothesis testing in order to preserve the integrity of the experimental design. The sample used for hypothesis testing was thus comprised of 70 students: 35 students in the control group (20 girls, 15 boys), and 35 students in the treatment group (19 girls, 16 boys).

Program Implementation

The program "dose" in this implementation was 16 hours, which is within the recommended range of 12 to 18 hours specified for the Basic HIPP training. However the implementation was weaker than that recommended in the Help Increase the Peace Program Manual. The training was understaffed and training assistants incompletely prepared. Serious classroom management problems interfered with the quality of program delivery. The pilot site was under-resourced to give support to the program, though the administration was eager to have it. Teachers at the school were accommodating to the evaluator, but not involved in program delivery nor were they consistently modeling the skills and values being taught by the program.

Findings

At pretest (before the HIPP workshops) there was no significant difference between control and treatment groups on any of the process or outcome variables. Statistically significant results ($p < .05$) are as follows:

- At posttest (after the HIPP workshops) boys in the treatment group reported significantly less victimization by peers than boys in the control group.

This means that boys in the HIP Program reported fewer incidents of being yelled at, hit or slapped, having things thrown at them, being

pushed, being asked to fight, or being threatened with harm by peers after the HIPP intervention than an equivalent group of boys (who had been randomly assigned to the control group instead of the HIP Program).

Victimization could not be legitimately measured for girls, because no scale could be derived (from items on the questionnaire related to victimization) that was statistically an acceptably good fit for the girls' data.

- At posttest boys in the HIPP workshops reported significantly greater self-efficacy for peaceful conflict resolution than boys in the control group.

This means that, after the HIPP intervention, boys in the HIP Program felt more confident that they could stay out of fights by using an apology, ignoring someone who was trying to make them mad, walking away, getting help from an adult and by generally knowing how to stay out of any fight that might start at school than an equivalent group of boys (who had been randomly assigned to the control group instead of the HIP Program).

The self-efficacy scale was valid and reliable for both girls and boys. The difference between treatment and control girls' self-efficacy after the HIPP intervention was not significant. (However, see the first "Within-group Before and After" finding bulleted below.) There was also a marginally significant ($p < .10$) interaction between treatment and sex indicating that the program may not be as effective for girls as for boys in the area of self-efficacy.

The implicit assumption that the greater boys' self-efficacy (a process variable) caused by their participation in HIPP led to (mediated) their lower victimization by peers (a behavioral outcome) was tested using multiple regression analysis. The mediation effect was not statistically significant. However, the statistical power to detect mediation of a behavioral outcome by a process variable in this study was limited by the small number of boys in an already fairly small sample. No other statistically significant effects were found in multiple regression analyses that compared treatment and control group students. Within-group Before and After Findings

Within-group findings, which compare before and after measures for each group separately, aid understanding of the experimental findings (those that compare treatment and control groups). Within group findings of interest are as follows [†]:

- Although there was no significant difference between treatment and control girls'

self-efficacy at posttest, girls in the control group showed a statistically significant decrease in self-efficacy from pretest to posttest while girls in the treatment group showed no significant change in self-efficacy.

- The victimization of boys in the HIPP trainings remained at the same level from pretest to posttest while the victimization of boys in the control group increased significantly from pretest to posttest leading to the significantly better outcome for treatment group boys as compared to control group boys at posttest.
- A non-significant pretest to posttest increase in self-efficacy for boys in the treatment group coupled with a non-significant decrease in self-efficacy for boys in the control group underlies the significant difference found between the two groups at posttest.
- In contrast with previous anecdotal reports that the intervention improves students' attitudes toward interpersonal conflict, the HIPP program caused no significant pretest to posttest changes in attitudes of girls or boys as measured in this evaluation. The control group students also showed no significant change in attitudes.
- The finding of no significant difference between treatment and control groups on curriculum knowledge at posttest is the result of both groups increasing their knowledge by the same amount (7%) from pretest to posttest.
- There were two within-group findings that were not in the expected direction. Control boys knowledge increased significantly from pretest to posttest and control girls' self-reported physical aggression decreased significantly pretest to posttest.
- Physical aggression for treatment and control groups did not show any significant change pretest to posttest when girls and boys data was pooled.

† The traditional significance level of $p < .05$ is used unless otherwise indicated.

Limitations, Generalizability and Recommendations

Limitations on Study Ability to Detect Outcomes

The size of the sample was small and this limited the power of the study to find statistically significant program effects. This makes the fact that some statistically significant changes were found, including a behavioral outcome in the expected direction (lower victimization for boys in the HIPP workshops compared with boys in the control group at posttest), all the more remarkable.

Measurement reliability for boys' victimization, self-efficacy and girls' non-violent choices was in a generally accepted range for social science research (Cronbach alphas in the .70s and .80s). Therefore the outcomes found using these measurement scales (i.e. positive effect on boys victimization and boys self efficacy, and no change on girls' attitudes) have reasonable credibility. However, the reliability of the measures for physical aggression and boys' attitudes supporting aggression was undesirably low (Cronbach alphas in the low .60s) and the knowledge test, which is under development, had a Cronbach alpha only in the low .50s. The lower reliability of these three measures means that "noise" in the measurement may have contributed to the lack of significant or consistent findings for physical aggression, boys attitudes, and knowledge.

Diffusion of treatment was evident between treatment and control group students who attended classes together: Students in the treatment group shared some program content information with control students. This may have diluted between group differences (i.e. lowered effect sizes).

Limitations Affecting Internal and External Validity of Outcomes

As a true experiment the design rules out most alternative explanations for observed outcomes other than the HIP Program itself. Some limitations remain, however. The only outcome measures used were self-report and thus susceptible to possible bias, especially in the direction of social desirability or under-reporting of victimization. Due to the experimental design, only bias that differentially affected treatment and control groups would have affected results. It is possible that students in the HIPP program might have wanted to answer in a socially desirable way to make the program look good. However, there is no evidence that this occurred and given the difficulties of implementation it seems unlikely. On the other hand, it seems more probable that diffusion of treatment coupled with compensatory effort by control students (trying harder to make up for not being in the group chosen to receive HIPP) accounted for an increase in knowledge test scores for control group students as great as the increase in knowledge test scores by treatment group students. Compensatory effort, especially by girls, to be or appear "good" in spite of not having been assigned to the treatment group seems the most likely explanation of the control group girls reporting less physical aggression at posttest than at pretest while the physical aggression of treatment girls and boys and control boys showed no significant change from pretest to posttest.

The regression models used in the analysis were based on the investigators' judgment (using previous research as a guide) and information available. Some factors

that influence students' aggressive behavior (for example gang membership or level of violence in the family) were not part of the model. A more complete model could explain more of the variance in student behavior. However, random assignment of students to treatment and control groups is expected to keep any unmeasured factors from biasing results.

The most serious limitations on generalizability to keep in mind are that the program was evaluated only for African American students in one middle school and there has to date been no follow-up to see if effects endure. Effect sizes were modest (5% lower victimization for treatment boys than control boys and 13% higher self-efficacy for staying out of fights for treatment boys compared control boys at posttest holding pretest scores constant.)

Applicability of Findings

The HIP Program appears to be effective in holding the line on victimization by peers of African American 6th grade boys in a challenged urban middle school and increasing their self-confidence in being able to stay out of fights. This makes it a promising program for these situations. Whether or not HIPP can have comparable results in other settings with other populations remains to be tested. The program appears to be more effective for boys than for girls. The fact that results for African American boys were accomplished in a challenged setting without a particularly strong implementation makes it more likely that the program can prove successful in other similar settings, as long as the trainer is willing to persist in modeling the program despite disruptions. Stronger implementation might improve effect sizes.

External Program Evaluation of HIPP

Report submitted to AFSC by Copper M. Coggins, Ph. D.

May, 2005