

**Collaborative Development of a Theory-Based Student
Assessment for a Violence Prevention Program Evaluation**

Norm Constantine and Kathleen Curry
School and Community Health Research Group
WestEd, Oakland, CA

Paper presented at the
American Evaluation Association Annual Conference
Chicago, IL, November, 1998

Acknowledgement. Our special thanks to the evaluation workgroup for their collaboration on the development and implementation of this study:

David Barret, Oak Manor Elementary School

Armand Brint, Mendocino County Office of Alcohol and Other Drug Programs

Norm Constantine, WestEd

Kathleen Curry, WestEd

David Delgado, Nokomis Elementary School

Melissa Dunken, Mendocino County Youth Project

Katie Koerper, Hopland Elementary School

Valerie Lawe, Mendocino County Office of Alcohol and Other Drug Programs

David Lilker, Potter Valley Middle School

Suzy Miller, Mendocino County Office of Alcohol and Other Drug Programs

Anne Oliver, Mendocino County Youth Project

Tina Tyson, Potter Valley Middle School

Marian Venaas, Nokomis Elementary School

Ned Walsh, Mendocino County Office of Alcohol and Other Drug Programs

Karin Wandrei, Mendocino County Youth Project

The Safe Haven Violence Prevention Program was funded by the US Department of Health and Human Services, Center for Substance Abuse Prevention, *Federal grant #1 HD 1 SPO 6368* to the Mendocino County Office of Alcohol and Other Drug Programs, Project Director Ned Walsh. The Safe Haven Violence Prevention Program evaluation was funded by a subcontract from the Mendocino County Office of Alcohol and Other Drug Programs to WestEd, Principal Evaluator Norm Constantine.

For further information or for copies of the assessment instruments, please contact:

Norm Constantine, Ph.D.

Director, School and Community Health Research Group

WestEd, Oakland City Center

500 12th Street, Suite 340

Oakland, CA 94607

Phone: (510) 302-4345 FAX: (510) 302-4269

E-mail: nconsta@WestEd.org

<http://www.WestEd.org/schrg>

Kathleen Curry, M.A.

Research Associate, School and Community Health Research Group

WestEd, Oakland City Center

500 12th Street, Suite 340

Oakland, CA 94607

Phone: (510) 302-4217 FAX: (510) 302-4269

E-mail: kcurry@WestEd.org

<http://www.WestEd.org/schrg>

Summary

This paper describes a fast-track process of collaboratively developing a program theory-based assessment instrument during the fourth year of a five year local implementation of a CSAP violence prevention demonstration program. The program consists of several interrelated components delivered in grades four through eight across three school districts: (1) a classroom curriculum, (2) paraprofessional-delivered individual counseling, (3) peer conflict-mediation training, and (4) community mobilization. The assessment instrument employed during the first three years of the program consisted of individual items matched to individual elements of the classroom curriculum scope and sequence, without the benefit of a theoretical framework. The data collected with this instrument had yielded counter-intuitive findings that program staff believed did not adequately represent the intended effects of the program. WestEd was selected at the beginning of the fourth year of the program to redesign the evaluation, including development of a new primary assessment instrument.

Because of time pressures to begin data collection for fourth year pre-tests, the instrument development timeline had to be condensed from the six to nine months typically required for this type of collaborative theory-based process to approximately three months. In addition, the three-hour driving time between the evaluation office and the program site limited the evaluation team to monthly on-site visits. To maintain the collaborative process under these limitations, at several stages we obtained written input in place of a more typical facilitated discussion and consensus building process, then organized this material within an evolving theoretical framework and presented it back to the group for feedback and revision. This allowed us to tap the collective local group wisdom and reinforce local ownership of the evaluation at every stage of the process within the constraints of the time and travel resources available.

Our process involved the following ten primary steps:

1. Assembly and orientation of an expanded local evaluation work group;
2. Elicitation of logic model elements (problems, goals, etc.) from each group member;
3. Development of a combined work-group visual logic model;
4. Development of a generic behavioral-change theoretical framework;
5. Preliminary organization of logic model elements into the theoretical framework;
6. Priority ratings of theoretically organized model elements by the evaluation work group;
7. Completion of program-specific theoretical framework;
8. Development of item pool by an item-writing panel of local staff and outside experts;
9. Piloting items with student focus groups; and
10. Reliability analyses of items and scales.

Logic Models and Theoretical Frameworks

Logic models are commonly used in evaluation research to visually depict a program's structure, organization, assumptions, and/or intended outcomes. Logic models range from the fully descriptive and atheoretical, or at least weakly theoretical, sometimes referred to as a "program model," to a true theoretical framework based on a prior body of evidence around a specific behavioral change or other theory. The later is relatively rare in evaluation and much more common in investigator-designed research studies. Among the reasons for this are (1) that evaluators often don't hear as strong a message about the importance of theory in their training as, say, a research psychologist or sociologist does, and (2) that it can be an exceedingly challenging task to discover after the fact the implicit theory that a particular program is based upon.

An example of a relatively atheoretical logic model would be the program models/maps used to succinctly describe a complex community initiative, sometimes referred to as a "theory of change" framework. These have been used to good benefit in consolidating stakeholders' inconsistent perceptions of their program and in translating a pile of overwhelming program proposals into much more manageable one pager visual depictions of the key components, so there is certainly value in this type of model. But the "theory" here usually consists of some set of generic components such as problems/causes/interventions/milestones/etc. An example of a middle ground approach (semi-theoretical?) might be one linking risk and protective factors, interventions, outcomes, and measures.

A true theoretical framework, on the other hand, typically is based on a specific existing theory of social behavior (or of whatever), or an eclectic amalgamation of two or more existing theories. A visual depiction of the framework is used to organize the program assumptions, strategies, and intended outcomes, and the theorized relationships among them. Fishbein's Theory of Reasoned Action (beliefs, attitudes, perceived social norms, behavioral intentions, etc.) is an example of the first; and Fisher's Information/Motivation/Behavioral Skills Theory is an example of the second. A modified combination of these two theories was adapted as the theoretical framework for this evaluation.

Local Evaluation Work Group

At the time we began work on the evaluation at the beginning of year four, the existing evaluation committee consisted of several program administration staff from the two primary collaborating agencies. We worked with the committee to carefully specify additional key stakeholders in the evaluation who might be able to make a commitment to attend monthly half-day meetings. We ended up forming an evaluation workgroup consisting of four administrative staff, two of the program paraprofessional counselors, four teachers representing the three participating districts and two school types

(elementary and middle school), and the two evaluators. The existing evaluation committee was renamed the executive committee and retained to meet monthly with the evaluators on administrative matters, however all evaluation design and interpretation work was assumed by the expanded evaluation workgroup. The expansion of this group allowed us to better tap the necessary local wisdom needed to design the most effective evaluation, and to achieve local ownership of the evaluation, thereby increasing evaluation utility (Patton, 1997). In addition, the evaluation work group was supplemented with intermittent meetings of larger groups of teachers, principals, and paraprofessional counselors.

From Logic Model to Theoretical Framework

The first step involved having each member of the evaluation work group provide written answers to four basic questions about the program:

1. What problems is the program trying to address?
2. What strategies is the program employing to address these problems?
3. What are the desired results?
4. If the program is successful, how will the strategies achieve these results?

The responses to these questions from the nine participating members of the evaluation work group were then transcribed and organized into nine visually depicted mini-logic-models. These consisted of the verbatim responses to the questions organized into four sequential boxes. From these, a mega-logic-model was created (see Figure 1). This involved integrating and summarizing the nine individual models into one model, while still maintaining the original responses from the nine participants, with wording sometimes condensed or slightly modified for clarity or consistency.

The evaluation team then reviewed the mega-logic-model and considered a variety of well-known health behavior change theoretical models for potential fit. We ultimately developed an integration of Fisher and Fisher's (1992, 1996) Information/Motivation/Behavioral Skills Theory and Fishbein's Theory of Reasoned Action (Fishbein and Middlestadt, 1989.) A generic program theoretical framework was then created for this adapted model (see Figure 2.) Finally, the individual elements contained in the mega-logic-model were reorganized and integrated into this theoretical framework in a series of lists.

The mega-logic-model and the generic program theoretical framework were then explained to and reviewed with the evaluation workgroup, and the list of program elements as organized by the theoretical framework was presented. The group was then asked to rate within each category their first, second, and third priority importance choices of the elements.

These ratings were then tabulated, and the lists were sorted by ratings and classified as highest importance elements, middle importance elements, and lowest importance elements (see Figure 3; highest rated items are in bold, lowest rated items are lined out, and summed ratings are in parentheses.) Finally, the most important elements were transferred onto the visual theoretical framework, and a selection of the middle-level important elements that appeared to the evaluation team to best fit within the framework were also included. What resulted was a concise and highly specific theoretical framework, closely fit to the group's implicit theoretical framework for the project while at the same time based in formal behavioral-change theory (see Figure 4.)

Item Development

A panel of local program staff and outside violence prevention experts developed a pool of approximately 200 potential items. The completed theoretical framework was used as a blueprint for the development and selection of items for the assessment instrument. This involve subsets of items to measure *information and knowledge, beliefs and attitudes, perceived social norms, behavioral skills: efficacy, behavioral skills: difficulty, behavioral intentions, behavioral outcomes, and behavioral impacts*. Item structure was informed by the structure employed by Fisher and Fisher (1996, pp. 116-117) in a study of changing AIDS risk behavior in a college student population. Most items employ three- or four-point options, including *true, sort of true, not very true, and not true at all; and always, most of the time, sometimes, and never*.

The initial items developed by the panel were then reviewed, modified, deleted, and/or integrated by a second and partly overlapping panel of experts. The evaluation team then constructed a pilot version of the instrument for on site pilot testing. The pilot test involved four focus groups of five students each representing the range of ages and level of functioning included in the program. As a result of this pilot test, items were further modified, deleted, and/or integrated, yielding a field test version of the instrument for grades 6 through 8, consisting of 73 items. A subset of scales and items were selected to form the elementary school version for grades 4 and 5, consisting of 47 items. Both versions of the instrument were also translated into Spanish by a professional translation team, and verified by several other Spanish speaking reviewers.

Field Test

Because of the compressed schedule, the instrument field-test had to be run concurrently with the collection of fourth year pre-test data. To maintain the integrity and usefulness of the pre-test assessment data, a decision was made that any modification to the instrument resulting from the field test data would be limited to deleting the least useful items. The field test versions of the instruments were purposely designed to have approximately a third more items than intended for the final version.

Reliability analyses. All 47 grades four through eight classrooms receiving the Safe Haven curriculum across the seven program schools participated, with 811 students returning signed parental permission forms and participating in the first assessment. Approximately half of the classrooms were randomly assigned to a delayed intervention control condition. Based on data from the first assessment, frequency distributions were calculated for all items. We then calculated scale scores and coefficient alphas for potential scales, and deleted items based on (1) lack of internal consistency with its potential scale, and/or (2) insufficient variance (i.e., almost all students responded the same way.) The resulting scales consist of three items each for the middle school (grades 6 through 8) Form B, and two or three items for the elementary school (grades 4 and 5) Form A (see Tables 1 through 8).

All data reported here are preliminary, pending completion of final data cleaning and verification procedures. The median coefficient alpha was .59 for Form A and .71 for Form B based on data from the first assessment. Because these were the same data used to select best contributing items and therefore capitalized on chance, cross-validation coefficient alphas were recalculated using data from the second and third assessments, yielding similar median alphas of .58 for the elementary grades and .70 for the middle school grades from assessment 2, and increased median alphas of .70 and .76 respectively from assessment 3. Overall, all scales showed acceptable or desirable coefficient alphas with the exception of two of the knowledge scales: violence definitions and effects of alcohol, tobacco, and other drug use. The contributing items to these scales suffered from insufficient variance in that almost all students answered most knowledge items correctly. The knowledge scales were retained, however, due to the importance of these items to the curriculum approach.

Reliability estimates based on a scale's stability over time (sometimes referred to as test-retest reliability) also were calculated by linking pre test to post test assessments. Post tests were administered approximately eight weeks after the pretest. The delayed-treatment experimental design employed provided for no intervention for Group A between assessments 2 and 3, and no intervention for Group B between assessments 1 and 2. Median test-retest Pearson correlation coefficients for the elementary school Form A were .49 and .53, and for the middle school Form B were .56 and .65 (see Tables 9 through 11).

Conclusions

We have demonstrated that it is possible to collaboratively construct the implied theory of an existing program, to frame it within a formal behavioral-change theory, and to develop a reliable assessment instrument directly tied to this theoretical framework. Construct validity analyses will next be conducted by analyzing the correlation patterns of the items and of the scales scores, and the theoretically predicted relationships among scale scores

and between scales scores and self-reported behaviors. Additional analyses will involve English and Spanish subsets of the sample, as well as other important subgroups as sample size permits.

References

- Fishbein, M. & Middlestandt, S.E. (1989). Using the theory of reasoned action as a framework for understanding and changing AIDS related behaviors. In V.M. Mays, G.W. Albee, & S.F. Schneider (eds.), *Primary Prevention of AIDS: Psychological Approaches* (pp. 93-110). London: Sage.
- Fisher, J.D. & Fisher, W.A. (1992). Changing AIDS risk behavior. *Psychological Bulletin*, *111*, 455-474.
- Fisher, J.D, Fisher, W.A, Misovich, S.J., Kimble, D.L., & Malloy, T.E. (1996). Changing AIDS risk behavior: effects of an intervention emphasizing AIDS risk reduction information, motivation, and behavioral skills in a college student population. *Health Psychology*, *15* (2), 114-123.
- Patton, M. Q. (1997). *Utilization-Focused Evaluation: The New Century Text*. Thousand Oaks, CA: Sage.

Figure 1. Program Logic Model

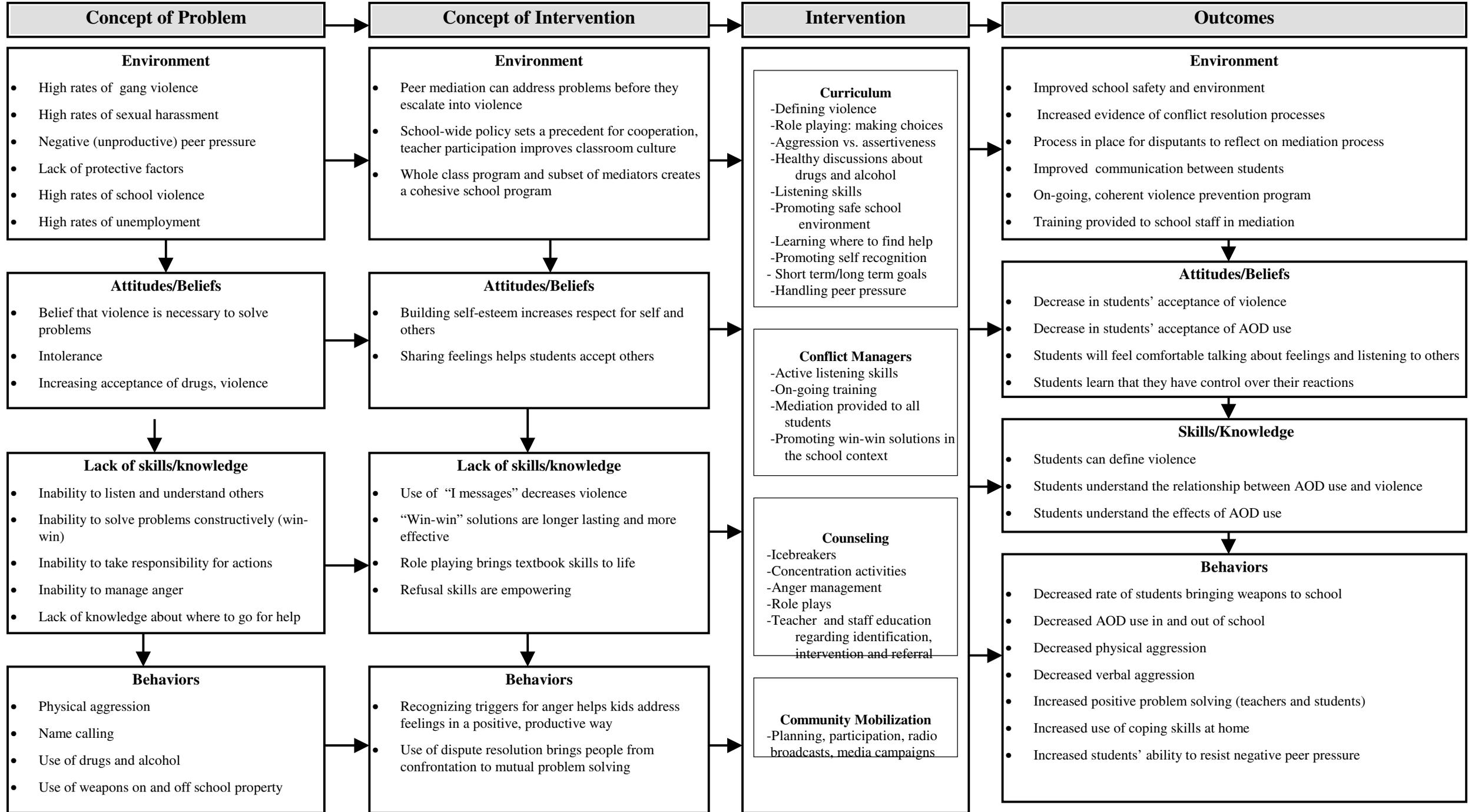


Figure 2. Program Theoretical Framework

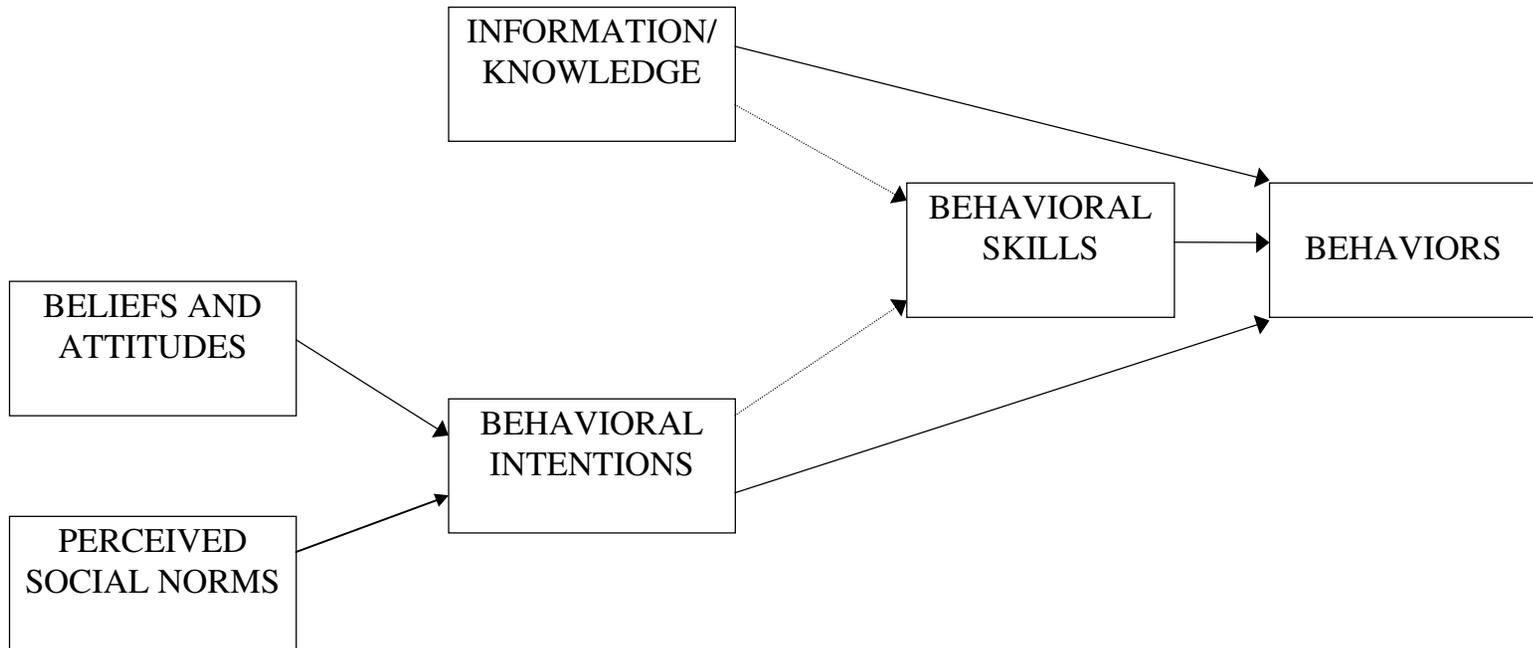


Figure 3. Program Theoretical Framework: Elicited and Rated Components

1. Information
 - **Relationship between AOD abuse and violence (13)**
 - **Where to find help (7)**
 - **Right to a safe school (7)**
 - Definition of violence (6)
 - Aggression versus assertiveness (6)
 - Win-win concepts (5)
 - Recognizing own strengths and limitations (5)
 - ~~Facts about drugs and alcohol (0)~~

2. Attitudes
 - **Respect for self (16)**
 - **Respect for others (14)**
 - Acceptance of others (9)
 - Recognition of self (9)

3. Perceived Social Norms
 - **Acceptance of violence (18)**
 - **Belief that violence is necessary to solve problems (12)**
 - Acceptance of drugs (9)
 - Intolerance (9)

4. Behavioral Intentions
 - **To take responsibility for actions (23)**
 - **To address anger in positive way (13)**
 - To listen to others (10)
 - ~~To help others (2)~~

5. Behavioral Skills
 - **Anger management (15+5)**
 - **How to listen/active listening (11)**
 - Conflict management (8)
 - Use of coping skills (7+)
 - Determining goals (4+2)
 - How to help others (3)
 - ~~Concentration (0)~~

6. Behavioral Outcomes

Increase:

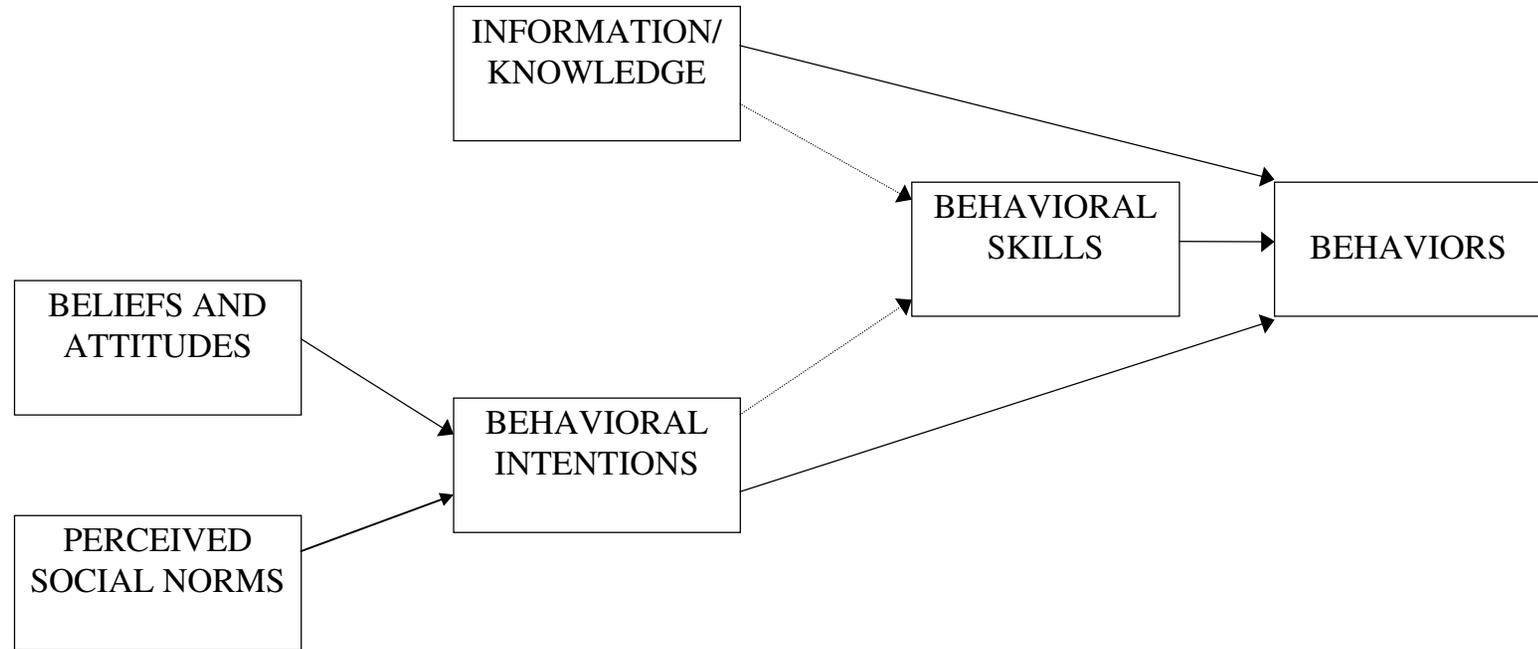
 - **Healthy/positive choices (12)**
 - **Positive problem solving (11)**
 - AOD drug use in school (4)
 - Acts of kindness and courtesy (3)
 - Taking time before reacting (3)
 - Use of “I” messages (2)
 - Participation in healthy activities (2)
 - Protective factors (unspecified) (2)
 - ~~Sharing feelings with others (0)~~
 - ~~Filling time doing positive things (0)~~

7. Behavior Impact

Decrease:

 - **Verbal aggression (13+3)**
 - **Physical aggression (13)**
 - Gang violence (3)
 - ~~Sexual harassment (1)~~
 - ~~Quarrels (1)~~
 - ~~Weapons to school (0)~~

Figure 4. Program Theoretical Framework with Measurable Elements



Attitudes

- Communication
- School environment
- Conflict management

Perceived social norms

- Alcohol
- Drugs
- Physical aggression
- Verbal aggression
- Communication

Information/Knowledge

- Definitions of violence
- AOD and violence
- Effects of AOD

Behavioral Intentions

- Temptation refusal
- Assertive communication
- Trusting communication
- Accepting help
- Initiating help
- Self-control

Behavioral Skills

- Temptation refusal
- Assertive communication
- Trusting communication
- Accepting help
- Initiating help
- Self-control

Behaviors

- Use of alcohol
- Use of drugs
- Physical aggression (2)
- Verbal aggression
- Communication

November 3, 1998

Table 1. Scale Score Coefficient Alphas from Assessment 1

Form A: Grades 4 and 5, Version 1

Number of Students = 410

Data from January, 1998 (Assessment 1)

SECTION/SCALE NAME	DESCRIPTION	ITEMS	ALPHA
Section B (Attitudes and beliefs)			
BLCOMM45	Communication	-B1, -B4, -B9	.49
BLSCEN45	School Environment	B2, B6, -B12	.59
BLCNFM45	Conflict Management	-B5, -B8, -B11	.53
Section C (Perceived Norms)			
PNALCL45	Alcohol	C1,C3	.59
PNPVAG45	Physical and Verbal Aggression	-C2, C4, C5	.53
PNCOMM45	Communication	-C7, -C8, -C9	.58
Section D (Knowledge)			
KNVLDF45	Violence definitions	-D1, D2, -D8	.06
KNDAEF45	Effects of AOD	D4,-D5, D11	.40
Section E-A (Perceived Self-Efficacy)			
SECMTR45	Trusting communication	E2, E6, E11	.66
SEHLAC45	Accepting help	E3, E9	.61
SEHLIN45	Initiating help	E4, E7, E12	.57
SESECO45	Self-control	E5, E14	.69
Section E-B (Behavioral Intentions)			
BICMTR45	Trusting communication	E2, E6, E11	.74
BIHLAC45	Accepting help	E3, E9	.49
BIHLIN45	Initiating help	E4, E7, E12	.60
BISECO45	Self-control	E5, E14	.58
Section F (Behaviors)			
BHUSAL45	Use of alcohol	F2, F8, F10	.74
BHPAG45A	Physical aggression - 1	F3, F7, F9	.74
BHVRAG45	Verbal aggression	F11, F12	.79
BHCOMM45	Communication	-F1, -F4, -F14	.59

November 3, 1998

Table 2. Scale Score Coefficient Alphas from Assessment 1

Form B: Grades 6, 7 and 8, Version 1

Number of Students = 401

Data from January, 1998 (Assessment 1)

SECTION/SCALE NAME	DESCRIPTION	ITEMS	ALPHA
Section B (Attitudes and beliefs)			
BLCOMM68	Communication	-B1, -B4, -B9	.61
BLSCEN68	School Environment	B2, B6, -B12	.66
BLCNFM68	Conflict Management	-B5, -B8, -B11	.61
Section C (Perceived Norms)			
PNALCL68	Alcohol	C1, C8, C17	.82
PNDRGS68	Drugs	C4, C6, C14	.83
PNPHAG68	Physical Aggression	C2, C7, C16	.60
PNVRAG68	Verbal Aggression	C5, C9, C11	.75
PNCOMM68	Communication	-C13, -C15, -C18	.76
Section D (Knowledge)			
KNVLDF68	Violence Definitions	-D1, D2, -D10	.16
KNDAVL68	AOD and Violence	D5, D11, D13	.63
KNDADR68	Effects of AOD	-D7, -D15, -D17	.69
Section E-A (Perceived Self-Efficacy)			
SERFSL68	Temptation Refusal	E3A, E10A, E21A	.85
SECMAS68	Assertive Communication	E12, E17, E19	.71
SECMTR68	Trusting Communication	E5, E8, E14	.83
SEHLAC68	Accepting Help	E4, E11, E15	.71
SEHLIN68	Initiating Help	E6, E18, E20	.64
SESECO68	Self-Control	E7, E16, E22	.77
Section E-B (Behavioral Intentions)			
BIRFSL68	Temptation Refusal	E3, E10, E21	.85
BICMAS68	AssertiveCommunication	E12, E17, E19	.71
BICMTR68	Trusting Communication	E5, E8, E14	.83
BIHLAC68	Accepting Help	E4, E11, E15	.73
BIHLIN68	Initiating Help	E6, E18, E20	.67
BISECO68	Self-control	E7, E16, E22	.78

Section F (Behaviors)			
BHUSAL68	Use of Alcohol	F4, F11, F13	.83
BHUSDR68	Use of Drugs	F15, F18, F19	.55
BHPAG68A	Physical Aggression - 1	F1, F5, F10	.76
BHPAG68B	Physical Aggression - 2	F12, F20, F24	.63
BHVRAG68	Verbal Aggression	F14, F22, F23	.84
BHCOMM68	Communication	-F3, -F6, -F26	.67

Tables 3 – 10 available upon request.

(additional scale score coefficient alphas for assessments 2, 3, and 4;
and test-retest reliabilities)

November 3, 1998

Table 11. Median Reliability Estimates

	Form A median alpha	Form A median test- retest	Form B median alpha	Form B median test- retest
Assessment 1	.59		.71	
Assessment 2	.58		.70	
Assessment 3	.70		.76	
Assessment 1-2*		.49		.56
Assessment 2-3*		.53		.65

* no intervention group only

- Form A – elementary school (grades 4 and 5) version
- Form B – middle school (grades 6 through 8) version