

Section C: Evaluating the Effectiveness of the Enhanced Men's Non-Violence Program

Background

One component of a coordinated community response to domestic violence is the provision of rehabilitation services to offenders. The Men's Non-Violence Program in Duluth uses an educational group format based upon a curriculum developed by staff at the DAIP, which is used throughout the country (Pence & Paymar, 1993). Each offender referred to the program attends educational group sessions offered at human service agencies, including the Domestic Abuse Intervention Project. Until recently, these groups have operated in relative isolation to the broader community. Beginning in late 1996, the DAIP provided an alternative program in which men's non-violence groups integrated community meetings with the ongoing activities of small groups of men led by a facilitator.

The purpose of this study is to determine if the alternative program that integrated community meetings was an effective enhancement to ongoing program groups on several criteria. Part one examines recidivism data and critical issues relating to program implementation. Part two examines abuse, safety and well-being based upon the results of victim interviews. Part three discusses the overall findings.

Part One: Recidivism

Hypothesis and Research Question

- | | |
|-------------------|---|
| Hypothesis 4a. | Men who attended the Alternative Men's Nonviolence Program in 1997 will have lower rates of recidivism when compared to those who attended the Lengthened Standard Men's Nonviolence Program in 1997. |
| Hypothesis 4b. | Men who attended the Alternative Men's Nonviolence Program during the first six months of 1998 will have lower rates of recidivism when compared to those who attended the Lengthened Standard Men's Nonviolence Program in 1998. |
| Hypothesis 4c. | Men who attended the alternative Men's Nonviolence Program in 1997 will have lower rates of recidivism when compared to those who attended the Standard Men's Nonviolence Program in 1996. |
| Hypothesis 4d. | Men who attended the Alternative Men's Nonviolence Program during the first six months of 1998 will have lower rates of recidivism when compared to those who attended the Standard Men's Nonviolence Program in 1996. |
| Research Question | What are critical issues related to the alternative men's non-violence program that should be considered in examining the results? |

Method

Population and Sample

The population for the study was the men entering the Men's Nonviolence Program under the auspices of the Domestic Abuse Intervention Project. The sample consisted of 515 men in the following groups:

- 195 - standard program (1996)
- 114 - lengthened standard (1997)
- 106 - alternative program (1997)
- 50 - lengthened standard (1998)
- 50 - alternative program (1998)

Tables 1 and 2 summarize the demographic data available on male offenders entering the DAIP programs and their female victims for each year of the study. Some 1996 offenders were dropped from the study because they were included in a pilot test of the alternative and lengthened program. Race and age were the only demographic variables available from agency databases.

Table 1

Demographic Data on Male Domestic Violence Offenders

Variable	Year			
	1994 (n=261)	1996 (n=217)	1997 (n=220)	1998 (n=100)
Age				
Under 21	13 (5%)	6 (3%)	9 (4%)	7 (7%)
21-30	116 (44%)	81 (37%)	93 (42%)	31 (31%)
31-40	82 (31%)	85 (39%)	75 (34%)	38 (38%)
41-50	35 (13%)	33 (15%)	38 (17%)	21 (21%)
Over 50	11 (4%)	12 (6%)	5 (2%)	2 (2%)
Unknown	4 (2%)	0 (0%)	0 (0%)	1 (1%)
Race				
American Indian	32 (12%)	35 (16%)	32 (15%)	17 (17%)
Asian American	1 (0%)	0 (0%)	0 (0%)	2 (2%)
African American	16 (6%)	15 (7%)	14 (6%)	8 (8%)
Hispanic	0 (0%)	3 (1%)	1 (1%)	2 (2%)
White	209 (80%)	159 (73%)	172 (78%)	71 (71%)
Other	3 (1%)	5 (2%)	1 (1%)	0 (0%)

Table 2Demographic Data on Female Domestic Violence Victims

Variable	Year							
	1994 (n=261)		1996 (n=217)		1997 (n=220)		1998 (n=100)	
Age								
Under 21	24	(9%)	20	(9%)	22	(10%)	7	(7%)
21-30	117	(45%)	75	(35%)	82	(37%)	36	(36%)
31-40	62	(24%)	77	(36%)	59	(27%)	30	(30%)
41-50	23	(9%)	30	(14%)	33	(15%)	13	(13%)
Over 50	10	(4%)	4	(2%)	6	(3%)	2	(2%)
Unknown	25	(10%)	11	(5%)	18	(8%)	12	(12%)
Race								
American Indian	23	(9%)	25	(12%)	33	(15%)	12	(12%)
Asian American	1	(0%)	2	(1%)	0	(0%)	3	(3%)
African American	4	(2%)	7	(3%)	5	(2%)	2	(2%)
Hispanic	0	(0%)	2	(1%)	0	(0%)	0	(0%)
White	227	(87%)	181	(83%)	175	(80%)	82	(82%)
Other	0	(0%)	0	(0%)	1	(1%)	0	(0%)
Unknown	6	(2%)	0	(0%)	6	(3%)	1	(1%)

Design

A post-test only control group design with randomized assignment was used to test the hypotheses comparing the alternative (AP) and the lengthened standard (LSP) program participants during 1997 and 1998. Men were randomly assigned to either AP or the LSP by intake staff using a computer-generated assignment list. A non-equivalent comparison group design was used to compare the 1997 and 1998 AP participants to the 1996 standard program (SP) participants. In addition, repeated follow-up measurements were used in both designs. Statistical procedures were used to determine whether or not demographic and programmatic variables had an impact on recidivism.

Information regarding the research question on issues related to the men's non-violence program was obtained in two ways. First, evaluation forms were completed at each community meeting, and a sample of four of these evaluations was examined to identify strengths and weaknesses of the meetings, along with other issues. The four meetings from which forms were analyzed were held in January, March, August, and October, 1997; eight to eighteen men completed forms at the four meetings.

Secondly, a questionnaire related to community meetings was distributed to the men in non-violence classes in March, 1998. This questionnaire (see Appendix) asked men to identify strengths of the meetings and to suggest ways to improve the community meetings. They were also asked about reasons for not attending community meetings.

Operational Definitions and Data Collection

Standard Program. The men's non-violence program that existed in 1996 included 27 one and one-half hour group sessions in which men examined their own abusive behavior and explored non-violent alternatives to this behavior. The power and control wheel developed by Pence and Paymar (1993) served as a basis for examining abusive behavior and violence. Men were encouraged to acknowledge ways in which they had used a variety of means to exert power and control, including using intimidation, emotional abuse, isolation, minimizing, children, male privilege, economic abuse, and coercion and threats. Men then explored and practiced non-violent alternatives to abusive behavior that focused on equality in relationships, including non-threatening behavior, respect, trust and support, honesty and accountability, responsible parenting, shared responsibility, economic partnership, and negotiation and fairness.

The process for addressing the topics and themes identified above was consistent. A three week process included 1) defining the theme, 2) examining abuse behavior, and 3) exploring and practicing non-abusive and non-controlling behavior. Video segments of behavior between partners and personal examples of abusive behavior were used to provide a basis for reflection and discussion. Men were exposed to specific behavioral skills they could use to resolve problems without violence or coercive control. Finally, men were asked to relive their own abusive behavior through detailed role plays and to attempt to resolve the original situations without abuse.

Lengthened Standard Program. The standard men's program was lengthened in 1997 to include 33 class sessions. Each session lasted 90 minutes for a total of 49.5 class hours beyond the initial 3 hour intake session. The curriculum for these classes, as described above, remained the same.

Alternative Program. An alternative men's non-violence program was initiated in 1997. This program again maintained the same curriculum for 27 weekly class sessions of 90 minutes each. In addition each alternative program participant was required to participate in four community meetings, each of which lasted about two and one-quarter hours. Alternative program participants therefore attended 40.5 hours of class and 9 hours of community meetings for a total of 49.5 hours beyond the initial 3 hour intake session.

The community meetings were intended to link men in the non-violence program to other men in the community with a focus on the general principles of change rather than the specific behavior of participants. In linking the broader community to DAIP, the program aimed to reduce the isolation of how domestic violence was viewed. The goals of the alternative program were to engage participants in a motivational , interactive group learning experience to support non-violence and to reinforce and complement the existing curriculum.

Each community meeting included a consistent format. The Men's Program Coordinator welcomed the participants, briefly explained the meeting goals and ground rules, and then introduced a guest presenter. The presenter told his story and/or led an activity for about 30 minutes and then responded to questions or comments from participants. Following a 10 minute break, participants moved into small groups for discussions. Questions were provided to groups to help focus the discussions. The entire group then came together toward the end of the meeting to share information from discussion in a session co-facilitated by the Men's Program Coordinator and the guest presenter. Finally, the presenter was thanked and participants completed an evaluation.

Recidivism. Rates were determined by collecting data from criminal justice data bases for St. Louis County and the Minnesota Bureau of Criminal Apprehension. For 1996 and 1997 male offenders, data were collected 6, 12 and 18 months after intake. Because of time constraints, 6 and 12 month follow-up data were available for men who entered the program during the first six months of 1998 only. Three levels of recidivism were identified with each level requiring a greater

level of documentation. Men were identified as recidivists when at the time of follow-up they fell into one or more of the following categories:

1. Investigated for a domestic violence related incident, but not charged.
2. Charged with a domestic violence related offense or been a respondent in an OFP hearing, but not convicted or had an OFP awarded.
3. Convicted of an offense related to another domestic violence related incident or been the respondent where an Order for Protection (OFP) was awarded.

Control variables. The following demographic and program variables were examined: age of victim and offender, race of victim and offender, court mandated, completion of groups, number of sessions attended, and batterer categorization assigned by probation officers and by Men's Nonviolence Program intake staff. Probation officers and intake staff assigned offenders to a category ranging from "one" to "four" as follows: 1) no battering history, 2) low level/ not escalating, 3) clear pattern of battering behavior/ likely to escalate and 4) high risk of serious harm.(See Appendix for Matrices).

Data Analysis

The percent of men who had recidivated was determined by identifying whether or not they fell into one or more of the recidivism categories of investigated, charged and convicted. Men who had recidivated one or more times were identified as recidivists. The number of men identified as recidivists in each group were compared using Chi-Square tests to determine statistical significance.

A Chi-Square, Fisher's Exact test or t-test was used to determine if control variables were different for offenders who recidivated and those who did not. Spearman correlation procedures were used to examine the relationship between recidivism and batterer categories assigned by probation officers and DAIP Men's Non-Violence Program staff. Forward stepwise logistic regression procedures were then used to determine which set of control variables, if any, discriminated between offenders who recidivated and those who did not.

Data from the community meeting evaluations and the questionnaire administered to men's groups in March, 1998 were analyzed to identify strengths and weaknesses of the community meetings. Frequencies and percentages of quantitative responses were calculated, and open-ended responses were qualitatively analyzed to identify the common themes running through various items.

Results

Program Data

Men in the AP attended fewer class sessions and were less likely to complete the program. Offenders entered the standard program in 1996 and, therefore, had a longer time to attend and complete the program by the time the study had ended. However, 1997 and 1998 offenders in the AP and LSP had a similar opportunity to attend and complete the program. Table 3 summarizes the data.

Table 3

Men's Non-Violence Program Data

Variable	Standard (n=195)	Lengthened (n=164)	Alternative (n=156)
Classes attended (mean)	19.0	18.7	15.2
Completed Program	106 (54%)	66 (40%)	53 (34%)
Court Mandated	151 (77%)	125 (76%)	120 (77%)

Recidivism

Tables 4 and 5 summarize the results of the comparisons between the alternative program (AP) and the lengthened standard program (LSP). During 1997, the AP offenders had lower rates of recidivism at 6 and 12 months, but the differences with LSP offenders were not statistically significant. By the 18 month follow-up, the results were in the opposite direction from the hypothesis with the LSP offenders having lower rates of recidivism. The 1998 LSP offenders also had lower rates of recidivism, but the differences were not statistically significant. The follow-up data for 1998 offenders may be somewhat unreliable because of delays in data entry by the criminal justice system. Because the 1998 groups were randomly assigned, this delay would be expected to affect the recidivism of both groups equally.

Table 4Comparison of 1997 Alternative Program to 1997 Lengthened Standard Program

% Recidivated

Follow-up	AMP (n=106)	LSP (n=114)	Chi Sq.	Prob.**
6 Months	28%	28%	5.62	0.29
12 Months	37%	38%	11.67	0.09
18 Months	46%	42%	13.09	--

** one-tail significance

Table 5Comparison of 1998 Alternative Program to 1998 Lengthened Standard Program

% Recidivated

Follow-up	AMP (n=50)	LSP (n=50)	Chi Sq.	Prob.**
6 Months	26%	14%	4.85	.15
12 Months	38%	28%	4.24	.19
18 Months	44%	32%	3.71	.30

** one-tail significance

Tables 6 and 7 summarize the results of comparisons between the AP and 1996 standard program (SP). No significant differences were found between the two groups for either 1997 or 1998. The rates were lower for the AP at the 6 and 12 month follow-up periods for both years, but at 18 months they were slightly higher for 1997.

Table 6Comparison of 1997 Alternative Program to 1996 Standard Program

% Recidivated				
Follow-up	AMP (106)	SP (195)	Chi Sq.	Prob.**
6 Months	28%	31%	4.88	.34
12 Months	37%	41%	11.27	.13
18 Months	46%	44%	13.09	.11

** one-tail significance

Table 7Comparison of 1998 Alternative Program to 1996 Standard Program

% Recidivated				
Follow-up	AMP (50)	SP (195)	Chi Sq.	Prob.**
6 Months	26%	31%	.43	.25
12 Months	38%	41%	.11	.37

** one-tail significance

Control Variables

Factor by factor analysis found that most of the demographic variables were not significantly related to recidivism at any of the follow-up periods. The one exception was the race of the victim. Offenders whose victims were white women were less likely to recidivate at 12 (p=.01) and 18 (p=.04) month follow-up periods.

Variables relating to the offender's involvement with the Men's Nonviolence Program were more likely to be related to recidivism than demographic variables. Men who were court mandated to the program were significantly more likely to have recidivated at 6 (p=.00), 12 (p=.00) and 18 (p=.00) month follow-up periods. Men who did not complete the program were also more likely to recidivate at 6 (p=.02), 12 (p=.00) and 18 (p=.00) month follow-up periods. At the twelve month follow-up period, men who attended fewer group sessions were more likely to recidivate (p=.03). This approached significance at 6 (p=.06) and 18 (p=.06) month follow-up periods. The type of program attended was related to recidivism at any of the follow-up periods.

The lower the batterer category assigned by probation officers, the less likely the offender was to have recidivated at 6 (p=.01), 12 (p=.02) and 18 (.02) month follow-up periods. At the eighteen month follow-up period, 10 of the 28 (36%) category “1” offenders recidivated, 23 of the 51 category “2” (45%) offenders recidivated, 28 of the 44 (64%) category “3” offenders recidivated and 2 of the 2 (100%) category “4” offenders had recidivated.

Batterer categories assigned by intake staff of the Men’s Nonviolence Program were only significantly related to recidivism at the 12 month follow-up period (p=.05). At the twelve month follow-up period, 7 of the 17 (41%) category “1” offenders had recidivated, 34 of the 110 (31%) category “2” offenders had recidivated, 53 of the 146 (36%) category “3” offenders had recidivated and 20 of the 35 (57%) category “4” offenders had recidivated. Batterer categories were not included in the logistic regression because of the large number of cases with missing data.

Two variables remained in the final model when forward stepwise logistic regression procedures included control variables and the alternative-standard program as an independent variable at the 18 month follow-up(See Table 8). Offenders who were not court mandated and who completed the programs were less likely to be recidivists. Using this model, recidivism could be predicted accurately in 46% of the cases, while no recidivism could be predicted in 72% of the cases. Overall recidivism could be accurately predicted in 60% of the cases based upon this model. The Chi-Square for the model was 13.65 with 2 degrees of freedom, which was significant at the .001 level. Variables not predicting recidivism were: age of the victim, age of the offender, race of the offender, race of the victim, number of group sessions attended and the type of program. At the 6 month follow-up period there was a significant relationship between age and recidivism that was not present at later follow-up periods. Younger women were less likely to have been victimized by offenders who later recidivated at 6 months, while older offenders were less likely to recidivate at 6 months. No variables remained in the final model when the alternative-lengthened standard programs were included as the independent variable.

Table 8

Logistic Regression: Control Variables and Alternative-Standard Program at 18 months

Variables	Recidivism at 18 months		
	Beta	Standard Error	Prob.
Court Mandated	-.99	.30	.00
Program Completion	-.58	.24	.02

Critical Issues

Results related to the research question provided important information on issues related to the alternative men’s program. The review of four sets of evaluation forms completed at the conclusion of community meetings throughout 1997 indicated that both the speakers and the small group sessions were rated highly at each community meeting throughout the year. Respondents often reported that they could relate to the speaker’s experiences and that his sharing set a tone for the rest of the meeting. Similarly, listening to the speaker and the small group discussions were most often mentioned as the best part of the meetings. The questions provided to the small groups

were viewed as helpful in the small group discussion. A common theme that ran through responses for all four meetings was the value of openly sharing emotions and feelings on a personal level, to be able to tell one's story and to listen to others in a safe context. In the words of one respondent, it was "OK to let down the wall and talk to one another."

While respondents were generally satisfied with the meetings as they were, some ideas for improvement were offered. Several respondents hoped that more men would attend, including men of different ages. Several respondents also indicated that there was not sufficient time to discuss issues in the small groups. Some respondents also commented on the need to improve the physical arrangements (more comfortable chairs, less background noise from other rooms) and the refreshments (did not like the coffee or cookies).

Responses from the Community Meetings Questionnaire that was completed by 34 men in the weekly non-violence small groups provided additional information about men's perceptions about the community meetings. Almost 60% of these men had been attending DAIP classes for four months or less, and 12 men (35%) had yet to attend a community meeting, even though only men who had community meetings as part of the program were asked to complete the questionnaire. Of the 21 men completing an item rating the value of the community meetings, 4 (19%) men rated the meetings as "very valuable," 8 (38%) rated the meetings as "mostly valuable," 8 (38%) rated the meetings as "somewhat valuable," and 1 (5%) rated the meetings as "not at all valuable." The most valuable aspects of the meetings cited by respondents were small group discussions (6 responses), listening to speakers (5 responses), and connecting or talking with other men in similar circumstances (5 responses). Ideas for improving the community meetings included more small group time (2 responses), more speakers, more stories (2 responses), and shorten the meeting (2 responses).

Questionnaire responses also provided information on reasons for participants missing community meetings. Of sixteen men who had missed at least one community meetings, nine respondents cited work conflicts and two cited not being notified about the meeting as the reason for missing. Of the 20 men who reported missing the February, 1998 meeting, six respondents cited work conflicts and two cited not being notified of the meeting as the reason for missing. The respondents ideas for making sure that men could attend community meetings regularly included 1) doing more effective advertising (five responses), 2) meeting at different hours during the day (four responses), 3) meeting on a day other than Monday (two responses), and 4) making the meetings mandatory (two responses).

Discussion

Conclusions

The data did not support the hypotheses that men in the alternative program would have lower rates of recidivism. In general, there were not statistically significant differences between the AP and other groups, although 18 month follow-up rates were higher for the AP.

In a factor by factor analysis, several other variables other than program type were related to recidivism. The following factors were related to lower recidivism rates: white victims, not being court mandated, completing the program, attending more sessions and being assigned a lower batterer category.

When all the variables were controlled for in a logistic regression, the only significant variables were that men who were not court mandated and men who completed the program were less likely to recidivate when the AP was compared to the SP. No variables were significant when the AP was compared to the LSP.

The data suggest that the AP may have had higher recidivism rates because they attended fewer sessions and were less likely to complete the program. Surveys of program participants indicated that some men reported not attending the community meetings because of scheduling conflicts or lack of information about them.

Limitations

The comparison between the AP and the SP did not use an experimental design. Variables other than the intervention may have influenced the results. The comparison between the AP and the LSP did use an experimental design. However, it appears that the AP was not as fully implemented as the LSP because of lower attendance and completion rates. It is not possible to determine whether or not there would have been significant differences between the two programs, if they both had been fully implemented. The recidivism data may not be reliable for 1998 because of delays in data entry into criminal justice data bases.

Part Two: Abuse, Women's Safety and Women's Well being

Hypotheses

- Hypothesis 4e Men who attended the Alternative Men's Nonviolence Program in 1997 will have lower rates of abusive behavior when compared to those who attended the Lengthened Standard Men's Nonviolence Program in 1997.
- Hypothesis 4f. Men who attended the Alternative Men's Nonviolence Program in 1997 will have higher ratings of victim safety when compared to those who attended the Lengthened Standard Men's Nonviolence Program in 1997.
- Hypothesis 4g. Men who attended the Alternative Men's Nonviolence Program in 1997 will have higher ratings of victim well-being when compared to those who attended the Lengthened Standard Men's Nonviolence Program in 1997.
- Hypothesis 4h. Men who attended the alternative Men's Nonviolence Program in 1997 will have lower rates of abusive behavior when compared to those who attended the Standard Men's Nonviolence Program in 1996.
- Hypothesis 4i. Men who attended the alternative Men's Nonviolence Program in 1997 will have higher ratings of victim safety when compared to those who attended the Standard Men's Nonviolence Program in 1996.
- Hypothesis 4j. Men who attended the alternative Men's Nonviolence Program in 1997 will have higher ratings of victim well-being when compared to those who attended the Standard Men's Nonviolence Program in 1996.

Method

Population and Sample

The population studied were the female victims of male domestic violence offenders who entered the Men's Non-Violence program during 1996 and 1997. The sample included all women who agreed to be interviewed eighteen month's after the offender's intake appointment. Women whose partners entered the program in 1998 were not interviewed because of time limitations.

Unfortunately, many of the women could not be reached for interviews. This was primarily due to being unable to locate them. Of the 217 women from 1996, 74% could not be located, 10% declined to be interviewed and 16% were interviewed. Of the 220 women from 1997, 76% could not be located for the interview, 11% declined to be interviewed and 13% were actually interviewed. The partners or former partners of the 1996 women (n=34) all attended the standard program. In 1997, 12 of the respondents had partners or former partners assigned to the lengthened standard program and 16 of the respondents had partners or former partners assigned to the alternative program. Table 9 summarizes the descriptive data available for victims from each year.

Table 9Demographic Data from Victim Interviews

Variable	Year		Year	
	1996 (n=34)		1997 (n=28)	
Race				
American Indian	3	(9%)	4	(14%)
African American	0		1	(4%)
White	30	(88%)	23	(82%)
Asian American	0		0	(0%)
Hispanic/ Latina	1	(3%)	0	
Age at time of interview				
18-25	7	(21%)	6	(21%)
26-35	14	(41%)	12	(43%)
36-49	12	(35%)	8	(29%)
50-65	1	(3%)	2	(7%)
Over 65	0		0	
Income at time of incident				
Less than \$10,000	11	(33%)	2	(7%)
\$10,001 to \$20,000	6	(18%)	9	(32%)
\$20,001 to \$30,000	5	(15%)	6	(21%)
\$30,001 to \$40,000	6	(18%)	7	(25%)
\$40,001 to \$50,000	3	(9%)	2	(7%)
More than \$50,000	2	(6%)	2	(7%)
Employed at time of incident				
Full-time	15	(44%)	2	(43%)
Part-time	7	(21%)	6	(21%)
Not at all	12	(35%)	10	(36%)
Employed at time of interview				
Full-time	14	(41%)	14	(52%)
Part-time	8	(24%)	6	(22%)
Not at all	12	(35%)	7	(26%)
Has children				
Yes	26	(77%)	23	(82%)
No	8	(23%)	5	(18%)

Design

As noted in Part One, a post-test only control group design with randomized assignment was used to compare the alternative (AP) and the lengthened standard (LSP) program participants during 1997. A non-equivalent comparison group design was used to compare the 1997 AP participants to the 1996 standard program (SP) participants. The women who had been the victims of these men when they entered the programs were contacted for 18 month follow-up interviews. Statistical procedures were used to determine whether or not demographic and programmatic variables were related to victim reports of abuse, safety and well-being.

Operational Definitions and Data Collection

Women were initially contacted by mail to request their participation in the study months 18 after the offender's intake appointment with the Men's Nonviolence Program. Women were asked to either return a signed consent form indicating their willingness to be interviewed or to return a form indicating that they did not wish to be contacted for an interview. A trained interviewer contacted the women and in most cases interviewed her by phone. If women did not respond, they received a second mailing. Follow-up phone calls were attempted when women did not respond to either of the first two inquiries.

The Survivor's Experiences Questionnaire used in the interviews is located in the Appendix. This in-depth interview took approximately one hour to complete. Areas covered in the interview included experiences with domestic violence services, offender risk factors, current and previous abuse on the part of the offender, and the women's current status in terms of their safety and well-being. Three main outcome variables were examined in this part of this study and are operationally defined below.

Frequency of Abusive Behavior. The frequency of abusive behavior as reported by victims was determined by using the Abusive Behavior Inventory (Shepard & Campbell, 1992) located in Part VII of the Survivor's Experiences Questionnaire (see Appendix). Four items were modified from the original inventory based upon a validity study of the instrument (Shepard & Campbell, 1992) and feedback from the evaluation team and program staff. Respondents were asked to rate retrospectively the frequency of physical and psychological abusive behavior during the three month period prior to the incident and their frequency during the most recent three months (approximately 16 to 18 months after the incident) on a scale from 1 to 5 (1=Never, 2=Rarely, 3=Occasionally, 4=Frequently, 5=Very Frequently). The Abusive Behavior Inventory (ABI) has physical and psychological abuse subscales which were examined separately for each respondent. Frequency ratings for each subscale item were summed and divided by the total number of subscale items to obtain subscale scores for physical and psychological abuse. Further detail about the ABI is provided in Section B of this report.

Women's safety. Item #159 on the Survivors' Experiences Questionnaire (see Appendix) was used to measure "women's safety." Women were asked: "How safe do you feel now—is it worse, better, or the same?"

Women's well being. Women rated themselves as "worse", "better", or "same" on items relating to their freedom to make decisions, how they feel emotionally, how they feel about their economic situation, how they feel physically, how they feel about social relationships, how they feel about their spiritual life, and how they feel about their relationship with their children (items #148- 156) on the Survivor's Experiences Questionnaire (see Appendix). A total well-being score was calculated for each woman by adding the number of items that received a "better" rating.

Programs. The Standard, Lengthened Standard and Alternative Programs are described in Part One of this section of the report.

Control variables. The following demographic and program variables were included in the statistical analysis: age of victim and offender, race of victim and offender, whether or not they were court mandated, whether or not they completed the program, number of sessions attended, and batter categorization by probation officers and by Men's Nonviolence Program staff. Probation officers and intake staff assigned offenders to a category ranging from "one" to "four" as follows: 1) no battering history, 2) low level/ not escalating, 3) clear pattern of battering behavior/ likely to escalate and 4) high risk of serious harm. See Appendix for Matrices.

Data Analysis

Frequency of Abusive Behavior. Physical and psychological subscale scores were determined as described above. A change score was calculated for each woman by subtracting the scores before intervention from the scores after intervention. The mean change scores for each program were compared for physical abuse, psychological abuse and total abuse. Analysis of Covariance tests were conducted to assess whether there were statistically significant differences in post-intervention scores between the AP and the LSP and between the AP and the SP controlling for differences in pre-intervention scores.

Women's safety. Ratings for each year of the study were summarized and compared using chi-square procedures.

Women's Well-Being. Women's ratings of individual items were compared for each of the program. Total well-being scores were also compared for an indication of overall well-being. T-tests were conducted to determine whether or not there were significant differences in mean scores when the programs were compared.

A Chi-Square, Fisher's Exact test or t-test was used to determine if control variables were related to abusive behavior, women's safety and women's well-being. Spearman correlation procedures were used to examine the relationship between batterer categories assigned by probation officers and DAIP Men's Non-Violence Program staff and women's safety. In order to examine the relationship between the categories and well-being and abusive behavior the categories were collapsed into two (i.e., 1 & 2; 3 & 4) and t-tests were conducted. Forward stepwise logistic procedures were then used to determine which set of control variables, if any, discriminated between women who reported feeling safer and those that did not. Forward stepwise multiple regression procedures were used to examine control variables in relation to abusive behavior and women's well being.

Results

Abusive Behavior

For each type of program, there were significant reductions in the frequency of physical, psychological and total abuse when a three month period before intervention was compared to a three period after intervention (Tables 10, 11 and 12).

Table 10

Comparison of Abusive Behavior Inventory Scores Before and After Intervention for the
Standard
1996 Men's Nonviolence Program

Variable (n=34) tail)	Before	After	t	Prob. (1
Psychological Abuse	3.01	1.82	6.59	.000*
Physical Abuse	2.36	1.27	7.08	.000*
Total Scores	5.38	3.10	7.29	.000*

*significant at the .01 level

Table 11

Comparison of Abusive Behavior Inventory Scores Before and After Intervention for the 1997
Lengthened Standard Men's Nonviolence Program

Variable (n=34)	Before	After	t	Prob. (1 tail)
Psychological Abuse	3.22	1.38	6.08	.000*
Physical Abuse	2.13	1.16	4.26	.001*
Total Scores	5.34	2.54	5.62	.000*

*significant at the .01 level

Table 12

Comparison of Abusive Behavior Inventory Scores Before and Intervention for the 1997
Alternative Men's Nonviolence Program

Variable (n=34) tail)	Before	After	t	Prob. (1
Psychological Abuse	3.35	1.99	5.01	.000*
Physical Abuse	2.23	1.50	3.35	.002*
Total Scores	5.57	3.49	4.91	.000*

*significant at the .01 level

Tables 13 and 14 compare mean change scores between the alternative program and the standard and lengthened standard programs. In general, the SP and LSP had greater decreases in abusive behavior than did the AP.

Table 13

Comparison of Abusive Behavior Inventory Change Scores for 1) the 1996 Standard Nonviolence Program and 2) the 1997 Alternative Nonviolence Program

Variable	<i>1996 Standard program</i> (n=34)	<i>1997 Alternative Program</i> (n=16)
Psychological Abuse	-1.19	-1.36
Physical Abuse	-1.09	-.72
Total Scores	-2.28	-2.08

Table 14

Comparison of Abusive Behavior Inventory Change Scores for 1) the 1996 Standard Nonviolence Program and 2) the 1997 Alternative Nonviolence Program

Variable	<i>1997 Lengthened Standard program</i> (n=12)	<i>1997 Alternative Program</i> (n=16)
Psychological Abuse	-1.84	-1.36
Physical Abuse	-.97	-.72
Total Scores	-2.81	-2.08

Analysis of Covariance tests were conducted to assess whether there were statistically significant differences in post-intervention scores between the three types of groups, controlling for differences in pre-intervention scores. In both comparisons (between, [1] the *1996 Standard Men's Nonviolence Program* and the *1997 Alternative Nonviolence Program* and [2] *1997 Lengthened Standard Nonviolence Program* and the *1997 Alternative Nonviolence Program*), the main effect for "program" was *not* statistically significant in respect to psychological abuse, physical abuse or the combined psychological and physical abuse score. Tables 15 to 20 summarize this data.

Table 15

ANCOVA Analysis of Differences in Post-Intervention "Psychological Abuse" Scores--Controlling for the Effect of Differences in Pre-Intervention scores for 1) the *1996 Standard Men's Nonviolence Program* and 2) the *1997 Alternative Nonviolence Program*.

Variable	df	F	Prob
Pre-Intervention Score	1	10.073	.003
Program (<i>1996 Standard vs. 1997 Alternative</i>)	1	.000	.993
Error	47	---	---

Table 16

ANCOVA Analysis of Differences in Post-Intervention "Physical Abuse" Scores--Controlling for the Effect of Differences in Pre-Intervention Scores for 1) the *1996 Standard Men's Nonviolence Program* and 2) the *1997 Alternative Nonviolence Program*.

Variable	df	F	Prob
Pre-Intervention Score	1	10.310	.002
Program (<i>1996 Standard vs. 1997 Alternative</i>)	1	1.949	.169
Error	47	---	---

Table 17

ANCOVA Analysis of Differences in Post-Intervention “Composite” Scores--Controlling for the Effect of Differences in Pre-Intervention Scores for 1) 1996 Standard Men’s Nonviolence Program and 2) the 1997 Alternative Nonviolence Program.

Variable	df	F	Prob
Pre-Intervention Score	1	10.656	.002
Program (<i>1996 Standard vs. 1997 Alternative</i>)	1	.451	.505
Error	47	---	---

Table 18

ANCOVA Analysis of Differences in Post-Intervention “Psychological Abuse” scores--Controlling for the Effect of Differences in Pre-Intervention Scores for 1) the 1997 Lengthened Standard Men’s Nonviolence Program and 2) the 1997 Alternative Nonviolence Program.

Variable	df	F	Prob
Pre-Intervention Score	1	8.065	.009
Program (<i>1997 Lengthened Standard vs. 1997 Alternative</i>)	1	1.945	.175
Error	25	---	---

Table 19

ANCOVA Analysis of Differences in Post-Intervention “Physical Abuse” Scores--controlling for the Effect of Differences in Pre-Intervention Scores for 1) the 1997 Lengthened Standard Men’s Nonviolence Program and 2) the 1997 Alternative Nonviolence Program.

Variable	df	F	Prob
Pre-Intervention Score	1	6.916	.014
Program (<i>1996 Lengthened Standard vs. 1997 Alternative</i>)	1	1.446	.240
Error	25	---	---

Table 20

ANCOVA Analysis of Differences in Post-Intervention “Composite” scores--Controlling for the Effect of Differences in Pre-intervention scores--for 1) the 1997 Lengthened Standard Men’s Nonviolence Program and 2) the 1997 Alternative Nonviolence Program.

Variable	df	F	Prob
Pre-Intervention Score	1	9.703	.005
Program (<i>1997 Lengthened Standard vs. 1997 Alternative</i>)	1	1.846	.186
Error	25	---	---

Control variables. Factor by factor analysis found that the demographic variables (ie., victim age and race, offender age and race) were not significantly related to abusive behavior change scores. The number of sessions attended and program categorizations were also not significantly related to abusive behavior change scores. Women reported significantly greater reductions in physical abuse for those offenders who had completed the program when compared to than those who had not ($p=.04$). There were no significant differences for program completion and psychological abuse change scores and total abuse scores. Women reported greater reductions in physical abuse ($p=.01$), psychological abuse ($p=.01$) and total abuse ($p=.00$) when the offender had been court mandated to attend the program.

No variables predicted psychological abuse or the total abuse score when forward stepwise multiple regression procedures included the AP - LSP as an independent variable. A significant regression equation was found for physical abuse change scores [$F(2,25)=8.46$, $p=.002$], with a R square of .40. Younger women reported greater reductions in physical abuse and women reported greater reductions when the offender had attended more group sessions. Table 21 summarizes these findings. Variables not included in the model were victim race, offender race, offender age, court mandate, completion of groups and whether the offender had been in the alternative or lengthened-standard program.

Table 21

Stepwise Multiple Regression: Control Variables, Alternative-Lengthened Standard Program and Physical Abuse

Variables	Physical Abuse Score		
	Beta	Standard Error	Sign. t score
Victim Age	.52	.01	.00*
Attendance	-.48	.01	.01*

* sign. at .05 level or above

A significant regression equation was found for AP-SP as the independent variable and psychological abuse change scores as the dependent variable with one variable in the equation [F(1,48)= 15.08, p=.00] and a R square of .19. Women whose partners were court mandated reported greater reductions in psychological abuse. Two variables were in the model for the physical abuse change score [F(1,48)= 9.60, p=.00] with a R square of .18. Women whose partners were court mandated also reported greater reductions in physical abuse. The standard program was related to greater reductions in physical abuse being reported. For the total abuse score one variable remained in the model [F(1,48)=15.84] with a R square of .25. Again, women whose partners were court mandated reported greater reductions in total abuse. Table 22 summarizes these findings. Variables not included in any of the models were victim age, offender age, victim race, offender race, attendance and completion of groups.

Table 22

Stepwise Multiple Regression: Control Variables, Alternative-Standard Program and Abuse Scores

Psychological Abuse Score			
Variables	Beta	Standard Error	Sign. t score
Court Mandate	.49	.40	.00*
Physical Abuse			
Court Mandate	.46	.31	.00*
Type of Program	-.28	.24	.04*
Total Abuse			
Court Mandate	.50	.60	.00*

* significance of .05 level or above

Women's Safety

Women were asked: "How safe do you feel now--is it worse, better, or the same?" For purposes of Chi-Square analysis, the categories of "same and "worse" were collapsed into one category of "not better". Most women reported that it was "better" for all programs (See Tables 23 and 24). When the Chi-Square, Fisher's Exact Test was run, no significant relationship between the SP and AP variables was found (Exact Significance, one-sided =.312). No significant relationship between the LSP and AP variables was found either (Exact Significance, one-sided =.612).

Table 23Frequency of Women’s Safety Ratings for 1996 Standard or 1997 Alternative Programs

Variable	<i>1996 Standard Program</i> (n=34)	<i>1997 Alternative Program</i> (n=15)
Better	26 (76%)	12 (80%)
Same or Worse	7 (21%)	3 (20%)

Table 24Frequency Distributions of Women’s Safety Ratings according for the 1997 Lengthened Standard or 1997 Alternative Programs

Variable	<i>1997 Lengthened Standard Program</i> (n=12)	<i>1997 Alternative Program</i> (n=15)
Better	10 (83%)	12 (80%)
Same or Worse	2 (17%)	3 (20%)

Control variables. Factor by factor analysis found that none of the demographic or program related variables studied were significantly related to the women’s ratings of their safety.

One variable remained in the final model when forward stepwise logistic regression procedures included control variables and the alternative-lengthened standard program as an independent variable and safety as the dependent variable (See Table 25). Older women were more likely to report that their safety was “better”. Using this model, reports of “better” safety could be predicted accurately in 100% of the cases, while “not better” responses could be predicted in 40% of the cases. Overall, safety could be accurately predicted in 89% of the cases based upon this model. The Chi-Square for the model was 6.31 with 1 degree of freedom, which was significant at the .012 level. Variables not predicting recidivism were: age of the offender, race of the offender, race of the victim, number of group sessions attended, court mandate and the type of program. No variables predicted safety when the alternative-standard program were included in the model.

Table 25

Logistic Regression: Control Variables, Alternative-Lengthened Standard Program and Safety

Variables	Safety		
	Beta	Standard Error	Prob.
Victim Age	.17	.08	.04

*statistically significant at .05 or above

Women's Well Being

Women were asked a series of questions to rate their well-being in various areas of their life, comparing their current situation to prior to intervention (ie., better, same, worse). The categories of "same" and "worse" were collapsed into one category of "not better" for purposes of analysis. Table 26 summarizes well-being ratings by program attended by the offender. The majority of women gave "better" ratings in all areas of their lives with the exception of "economic situation".

Table 26

Comparison of Women's Well-Being Ratings for the 1996 Standard, 1997 Lengthened Standard, or 1997 Alternative Programs

Variable	<u>1996 Standard Program</u> (n=34)	<u>1997 Lengthened Standard Program</u> (n=12)	<u>1997 Alternative Program</u> (n=16)
Freedom to make decisions	85%* (29/5)	100% (12/0)	94% (15/1)
Emotionally	68% (23/11)	92% (11/1)	81% (13/3)
Economic situation	50% (17/17)	58% (7/5)	50% (8/8)
Physically	77% (26/8)	92% (11/1)	75% (12/4)
Social relationships	71% (24/10)	92% (11/1)	69% (11/5)
Spiritual life	65% (22/12)	83% (10/2)	81% (13/3)
How children are doing	78% (21/6)	80% (8/2)	85% (11/2)
Relationship with children	70% (19/8)	90% (9/1)	92% (12/1)
Parenting skills	78% (21/6)	80% (8/2)	92% (12/1)

*Per cent of "better" ratings followed by frequency breakdown of "better" and "not better" responses

Neither the differences between *1996 Standard* and *1997 Lengthened Standard*, nor those between *1996 Standard* and *1997 Alternative* were statistically significant: $t(38)=.295$, $p=.385$ (1-tail) for *1996 Standard* and *1997 Alternative*; and $t(21)=1.06$, $p=.152$ (1-tail) for *1997 Lengthened Standard* and *1997 Alternative*. Tables 27 and 28 summarize these findings.

Table 27

Test of Differences in Well-Being for the 1996 *Standard*, or 1997 *Alternative* Programs

Variable	<u>1996 SP</u>		<u>1997 AP</u>		<u>Fisher's Sig. (1-Sided)</u>
	better	not better	better	not better	
Freedom to make decisions	29	5	15	1	.365
Emotionally	23	11	13	3	.258
Economic situation	17	17	8	8	.619
Physically	26	8	12	4	.586
Social relationships	24	10	11	5	.572
Spiritual Life	22	12	13	3	.197
How children are doing	21	6	11	2	.479
Relationship with children	19	8	12	1	.123
Parenting skills	21	6	12	1	.254

Table 28

Test of Differences in Well-Being Scores for the 1997 *Lengthened Standard*, or 1997 *Alternative* Programs

Variable	<u>1996 SP</u>		<u>1997 AP</u>		<u>Fisher's Sig. (1-Sided)</u>
	better	not better	better	not better	
Freedom to make decisions	12	0	15	1	.571
Emotionally	11	1	13	3	.417
Economic situation	7	5	8	8	.479
Physically	11	1	12	4	.267
Social relationships	11	1	11	5	.160
Spiritual Life	10	2	13	3	.643
How children are doing	8	2	11	2	.596
Relationship with children	9	1	12	1	.692
Parenting skills	8	2	12	1	.398

A total well-being score was determined for each women by adding the number of items that received a “better” rating. Two overall well-being scores were determined; one that included items relating to children and one that did not. This was done because not all women had children and, therefore, could not respond to all the items.

There were a total of nine well-being items and on average women whose partners were in the SP and AP reported that seven of these were better than they had been prior to intervention. Women whose partners were in the LSP reported that eight out of nine areas were better. Neither the differences between *1996 Standard* and *1997 Lengthened Standard*, nor those between *1996 Standard* and *1997 Alternative* were statistically significant: $t(48) = .662$, $p = .256$ (1-tail) for *1996 Standard* and *1997 Alternative*; and $t(25) = 1.39$, $p = .089$ (1-tail) for *1997 Lengthened Standard*

and *1997 Alternative*. Tables 29 and 30 summarize these findings.

Table 29Comparison of Well-being Scores on Items 1-9 (Excluding those Respondents without Children)

<u>Type of Program</u>	<u>Mean</u>
<i>1996 Standard (n=27)</i>	6.59
<i>1997 Lengthened Standard (n=10)</i>	7.70
<i>1997 Alternative (n=13)</i>	6.85

Table 30Comparison of Well-being Scores on Items 1-6 (Excluding Children's Well-being Questions)

<u>Type of Program</u>	<u>Mean</u>
<i>1996 Standard (n=34)</i>	4.15
<i>1997 Lengthened Standard (n=12)</i>	5.17
<i>1997 Alternative (n=16)</i>	4.50

Control variables. Factor by factor analysis found that the demographic variables (ie., victim age and race, offender age and race) were not significantly related to well-being scores. Having attended more groups was related to higher well-being scores when the child related items were excluded ($p=.01$), as was having the offender court mandated ($p=.01$). Batterer categorizations and program completion were not found to be significantly related to well-being.

The court mandated variable predicted well-being (items relating to children/ parenting excluded) when forward stepwise multiple regression procedures included the AP-SP as an independent variable. See Table 31. The significant equation for the well being scores was $[F(1,48)= 5.40, p=.02]$, with a R square of .10. Women whose partners were court mandated to attend the program reported greater well-being than those who were not. Variables not included in this model were age of victim, age of offender, race of victim, race of offender, attendance, completed groups, and the type of program attended. No variables predicted well-being scores including when the AP-LSP variable was included.

Table 31Stepwise Multiple Regression: Control Variables and AP-SP Well-Being Scores

Variable	Beta	Standard Error	Significance of t score
Court Mandated	-.32	.65	.02*

* significant at the .05 level or above

Discussion

Conclusions

Abusive Behavior. The hypotheses that the AP would have lower rates of physical and psychological abuse reported by victims when compared to the SP and the LSP was not supported by the results of the study. There were significant reductions in abuse for all programs, but there were no statistically significant differences when the AP was compared to the other programs.

Variables relating to the offender's program participation were found to be related to recidivism when a factor by factor analysis was done. Women reported greater reductions in physical abuse when the offender had completed the program. They reported greater reductions in physical, psychological and total abuse when the offender had been court mandated to attend the program.

When regression procedures were used to examine the control variables and AP-LSP as an independent variable, younger women and women whose partners attended more sessions reported greater reductions in physical abuse. When the AP-SP comparison was examined in regression equations, women whose partners were court mandated reported greater reductions in psychological, physical abuse and total abuse. Women reported greater physical abuse when the offender had attended the SP.

Women's Safety. The hypothesis that the AP would have higher ratings of women safety compared to the SP and LSP were not supported by the results of the study. Most women reported that their safety was "better" at the 18 month follow-up interview for all programs. Factor by factor analysis found that none of the demographic or program related variables studied were significantly related to the women's ratings of their safety. Older women were more likely to report their safety was "better" when regression analysis included the control variables and the AP-LSP comparison. No variables predicted safety when the AP-SP were compared in a regression equation.

Women's Well-Being. The results did not support the hypotheses that the AP would have higher ratings women's well-being when compared to the SP and LSP. A majority of women reported that different areas of their lives were "better" with the exception of their "economic situation", but there were not significant differences between the AP and the other two programs. In a factor by factor analysis greater well-being was related to two variables; the offender having been court mandated to attend, and the offender having attended a greater number of sessions. Women whose partners were court mandated reported greater well-being when the AP-SP was examined in a regression analysis including the control variables. No variables predicted well-being when the AP-LSP were compared in a regression equation.

Limitations

The comparison between the AP and the SP did not use an experimental design. Variables other than the intervention may have influenced the results. The comparison between the AP and the LSP did use an experimental design in that offenders were randomly assigned to the programs. However, the inclusion of women to be interviewed was not random in that attempts were made to contact all women, but many could not be located. Also, it appears that the AP was not as fully implemented as the LSP because of lower attendance and completion rates. It is not possible to determine whether or not there would have been significant differences between the two programs, if they both had been fully implemented.

Another limitation relates to the reliability and validity of the outcome measures. The measures of safety and well-being have content and face validity, but no information is available on reliability and on construct and criterion validity. One study provides evidence of the reliability and validity for the measure of abusive behavior (Shepard & Campbell, 1992).

Part Three : Discussion of Overall Findings (Part One and Two)

The Alternative Men's Non-Violence Program was not more effective in reducing recidivism rates and abusive behavior or in increasing women's ratings of their safety and well-being when compared to the Standard Men's Non-Violence Program and the Lengthened Standard Program. There is some evidence that men in the AP did not do as well as those in the other two programs. This may be related to the lower attendance and completion rates of the men in the AP. Program completion and higher attendance rates were related to improved outcomes. In general, men in the AP responded that they had found the community meetings beneficial, but for a variety of reasons did not consistently attend them.

Whether or not the men were court mandated to attend the programs was the variable most often related to program outcomes. Men who were court mandated were more likely to recidivate, but their partners or former partners reported greater reductions in abusive behavior and greater well-being. As noted earlier in this report, volunteers may be less likely to come to the attention of the criminal justice system. Women, however, report that their behavior does not improve as much as those that are court mandated. Further study of this issue is needed. Another variable that warrants further study is the use of batterer categorizations to predict program outcomes. Factor by factor analysis indicated that men with lower batterer categories had lower rates of recidivism.

Demographic variables related to the offender were not found to be related to program outcomes. In a factor by factor analysis, offenders whose victims had been white women were less likely to recidivate. Age was a significant variable in regression analysis with younger women reporting greater reductions in abusive behavior and older women reported greater safety when the AP and LSP were included as independent variables.

The experimental design used to compare the AP-LSP programs strengthens confidence in the findings that the AP did not lead to improved program outcomes. The program may have had beneficial effects that were not captured in this study or which could not be fully realized because of the incomplete implementation of the program. Women in all three programs reported improved life circumstances after their partners had participated in the them.