DOCUMENT RESUME

ED 359 274 UD 028 809

AUTHOR	Zhang, Quahwu
TITLE	Social Psychological Consequences of Interpersonal Relations: A Confirmatory Approach to Testing Deutsch's Theory of Cooperation and Conflict Resolution.
INSTITUTION	Columbia Univ., New York, NY. Teachers Coll. International Center for Cooperation and Conflict Resolution.
SPONS AGENCY	Grant (W.T.) Foundation, New York, N.Y.; National Center for Research in Vocational Education, Berkeley, CA.
PUB DATE	92
NOTE	313p.; Doctoral dissertation, Columbia University. A concise version appears in "The Effects of Training in Cooperative Learning and Conflict Resolution in an Alternative High School"; see UD 028 807. For related reports, see UD 028 810-812 and UD 028 820.
PUB TYPE	Dissertations/Theses - Doctoral Dissertations (041) -- Reports - Evaluative/Feasibility (142)
EDRS PRICE	MF01/PC13 Plus Postage.
DESCRIPTORS	Blacks; *Conflict Resolution; *Cooperative Learning; Doctoral Dissertations; Economically Disadvantaged; Educational Theories; Group Dynamics; High Schools; *High School Students; Inner City; *Interpersonal Competence; Minority Group Children; Models; Nontraditional Education; *Program Effectiveness; *Sensitivity Training; Urban Schools; Urban Youth
IDENTIFIERS	*New York City Board of Education

ABSTRACT

This study tested M. Deutsch's theory of cooperation and conflict resolution using an intervention project at an inner city alternative high school in New York City. The study was designed to test the theory by confirmatory structural modeling and by evaluating the intervention. The procedure involved a pre- and post-test procedure administered during the intervention at three high school campuses. The pre-post-matched data set amounted to 346 students, mostly Black and Hispanic American, from economically disadvantaged backgrounds. The study indicates that positive changes in interpersonal relations, induced by constructive conflict resolution, lead to higher academic achievement, and that this effect is mediated by increased internal locus of control. In general, the results from theory testing show that constructive conflict resolution, and possibly effective group working, have a positive impact on a student's interpersonal relations and consequently on that student's self-esteem, locus of control, mental and physical health, and academic achievement. Included are 19 figures; 25 tables; 4 appendixes containing summaries of 14 scales used, additional tables and figures, a description of the intervention plan, and computer programs used for the modeling; and a bibliography of 302 items. (JB)

SOCIAL PSYCHOLOGICAL CONSEQUENCES OF INTERPERSONAL RELATIONS: A CONFIRMATORY APPROACH TO TESTING DEUTSCH'S THEORY OF COOPERATION AND CONFLICT RESOLUTION

BY

QUANWU ZHANG

INTERNATIONAL CENTER FOR COOPERATION AND CONFLICT RESOLUTION
MORTON DEUTSCH, DIRECTOR
TEACHERS COLLEGE, COLUMBIA UNIVERSITY

DECEMBER 1991

SOCIAL PSYCHOLOGICAL CONSEQUENCES OF INTERPERSONAL RELATIONS: A CONFIRMATORY APPROACH TO TESTING DEUTSCH'S THEORY OF COOPERATION AND CONFLICT RESOLUTION

By

Quanwu Zhang

Dissertation Committee

Professor Morton Deutsch Sponsor
Professor Harvey Hornstein
Professor David Krantz
Professor Eric Leifer
Professor Jane Monroe
Professor David Rindskopf

Submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy
under the Executive Committee of the Graduate
School of Arts and Sciences
Columbia University

1992

ABSTRACT

Social Psychological Consequences of Interpersonal Relations: A Confirmatory Approach to Testing Deutsch's Theory of Cooperation and Conflict Resolution

Quanwu Zhang

Numerous studies have accumulated testing Deutsch's theory of cooperation and conflict resolution since his initial study in 1948. However, it has only recently become possible to test the theory at an integrative level, such that the intricate relationships among a large number of variables can be investigated simultaneously. Through integrative testing, possible intervening processes and/or common causes among effects and outcomes may be explicitly identified. The confirmatory structural modeling approach developed by Joreskog (1973) provides a powerful tool for this type of theory testing. Utilizing an intervention project at an inner city alternative high school undertaken by Deutsch, the present study is designed both to test Deutsch's theory by confirmatory structural modeling and to evaluate the intervention.

Most previous field experiments were focused on cooperative learning. The social psychological consequences of constructive conflict resolution, particularly their implications in education, have rarely been systematically studied in a real school setting. This study indicates that positive changes in interpersonal relations, induced by

constructive conflict resolution, lead to higher academic achievement, and that this effect is mediated by increased internal locus of control.

In general, the results from our theory testing show that constructive conflict resolution, and possibly effective group working, have a positive impact on a student's interpersonal relations which are measured by his/her social support and victimization. Consequently, this impact leads to positive changes in the student's self-esteem, locus of control, mental and physical health, and academic achievement. In addition, some of these social psychological and educational changes may further enhance the student's ability to handle conflicts constructively and to solve problems more effectively when working with his/her group members.

The evaluation study has indicated that when cooperative learning is introduced together with conflict resolution training, students may be able to work most effectively in groups.

Issues regarding multiple structural modelings with a large number of variables in one study are discussed; a number of suggestions are made in the paper.

Contents

Figures

Tables

ACKNOWLEDGEMENTS

This research was supported by the William T. Grant Foundation and the National Center for Research in Vocational Education. Appreciation is extended to the students, teachers, and staff from the inner city alternative high school for their participation and cooperation in the process of this study.

I am especially grateful to my sponsor, Professor Morton Deutsch, for his unconditional support and inspiring advisement throughout my graduate life in this country. I am indebted to my dissertation committee members, Harvey Hornstein, David Krantz, Eric Leifer, Jane Monroe, and David Rindskopf for their advice and encouragement. I also wish to thank the research and training staff of the International Center for Cooperation and Conflict Resolution for their contribution to the intervention project. The excellent administrative support from Steve Turley was indispensable for the completion of this study. Special thanks are given to Frances LaFleur for the critical support I received from her.

Q.Z.

Chapter One

THEORETICAL OVERVIEW

Forty three years ago in his pioneering study (1948, 1949a, 1949b, 1968, 1973, 1985) of cooperation and competition, Deutsch empirically demonstrated that, by inducing either cooperation or competition, the interpersonal relations among subjects were significantly affected. Deutsch categorized these effects in terms of three basic social psychological processes which he described as substitutability (how one person's action can satisfy another person's needs), cathexis (positive or negative attitudes resulted from each other's action), and inducibility (the extent that one person can influence another by his or her action).

The study was conducted at MIT. Subjects from an introductory psychology course were randomly assigned to either a cooperative or a competitive grading system. In the cooperative grading system, all students within a given group received the same grades; their grades were determined by the level of the group's performance in comparison with other similar groups. In the competitve group, each student's contribution to the group performance was ranked in comparison with the contributions of other group members; and each student's grade was determined by his relative rank in his group.

The findings of the study can be best described in a 2x3 matrix arrangements (two conditions by three processes). The major hypotheses were supported by the results. By summarizing this original study and further developments, we intend to outline the main theme and statements of Deutsch's theory of cooperation and competition which has later evolved into the theory of cooperation and conflict resolution.

Deutsch's Theory of Cooperation and Competition

Two basic notions

1. Goal interdependence.

In the Lewinian approach, how people relate to each other in their interaction can be characterized in terms of the overlapping of their respective goal regions which Deutsch defined as "goal interdependence". There are two basic types of goal interdependence: promotive interdependence, where the goals are positively linked in such a way that the amount of the goal one person obtains is positively correlated with the amount that other people obtain; and contrient interdependence, where the goals are negatively linked in such a way that the amount of goal attainment of one person is negatively associated with the other's goal attainment.

Pure promotive or pure contrient interdependence rarely exists in real life. This can be seen clearly through combining type of goal interdependence with another basic idea of Deutsch's--types of action.

2. Basic types of actions

Deutsch characterized two types of actions by an individual: effective actions, which improve the actor's chances of obtaining his or her goal, and bungling actions, which worsen the actor's chances of obtaining his or her goal. When goal interdependence exists, the implications of their actions to the people involved can be rather complex. Deutsch used tennis as an example. In a doubles game, you and your partner are in a promotively interdependent relationship. When your partner allows an easy shot to get past him, you will have to extend yourself to prevent yourself from being harmed by the error. Thus the bungling action of your partner can not be a substitute for effective actions you had intended and the bungling will be cathected negatively. On the other hand, if your relationship is one of contrient interdependence, and the other bungles (as when your tennis opponent double-faults), your opponent's bungle will substitute for an effective action on your part, and it will be cathected positively or valued.

This illustration indicates that the same type of action of one party may contain quite different information to another depending upon their relationship in that given situation. Further, the psychological consequences of the action may either facilitate or jeopardize the existing relationship between the parties involved. For instance, an effective action from your game partner may be a substitute for your effective action and be valued positively on your part, thus enhancing the promotive interdependence between you two. While an effective action from your opponent will be perceived negatively, thus intensifing your desire to defeat him. It is in this context that Deutsch's crude law is relevant.

Deutsch's Crude Law of Social relations

Deutsch's crude law is simple and straight forward, therefore it is crude. But it portrays one of the fundamental phenomena of social relations which often tends to be ignored. The crude law says: the characteristic processes and effects elicited by a given type of social relationship also elicit that type of social relationship. A cooperative relationship facilitates each other's goal attainment, fosters positive attitudes toward each other, and induces mutual help and support, these in turn strengthen the cooperative relationship. While a competitive relationship creates tension among the people, causes hindrance to each other's action, brings negative feelings towards each other, these in turn escalate the competitiveness.

Deutsch's crude law subsumes numerous topics of social/ organizational psychology by underlining a common mechanism --reciprocal causation between interpersonal relations or social climate and intrapersonal processes. In terms of the applied research of social psychology, Deutsch's crude law can be used as a general guideline for social psychological interventions. The key point is to focus our intervention on the institutional changes which have hypothesized effects on the targeted psychological processes. Once the intended effects occur, the whole process of the intervention can be expected to enter into a positive reinforcement loop.

This is of course not to propose a "perpetual motion machine" for applied social psychology. However, if an intervention has been well received and internalized by the implementing institution, their self-generated programs can later function as Maxwell's demon which stablizes or reenergizes the ongoing changes.

The connection between cooperation vs competition and conflict resolution

The juxtaposition of cooperation vs. competition and the two types of goal interdependence in Deutsch's theory suggests that the former can be defined accordingly by the nature of the latter. In fact, they have been used interchangably in his papers. Thus if conflicts occur in a promotively interdependent relationship and are so perceived by the parties involved, the conflict resolution is more likely to take a constructive course. Since a cooperative process leads to the defining of conflicting interests as a mutual problem, it fosters the constructive resolution of conflicts by collaborative effort.

In his article (1987), Deutsch delineated the connection between cooperation vs. competition and types of conflict resolution: a constructive process of conflict resolution is similar to an effective, cooperative problem-solving while a destructive process of conflict resolution has many of the characteristics of a competitive process of social interaction. Thus these two basic types of conflict resolution are, in essence, the processes of cooperation or competition manifested in interpersonal relations. This is an important linkage in that the earlier studies of cooperation and competition can be largely applicable to the area of conflict resolution, which in fact has given rise to the integrated paradigm of cooperation and conflict resolution.

Cooperation and competition, however, are broader concepts than constructive or destructive conflict resolution, as the former obviously includes human interactions where no conflicts, at least no conflicting interests, are involved. When such a non-confrontational environment exists, the introducing of cooperative processes can be viewed as means of enhancing work effectiveness and productivity through improving communication, coordination and building mutual trust. In any group-working setting, therefore, the application of the theory of cooperation and competition can be conceived as having two main components: effective group working and constructive conflict resolution.

Main Propositions and Empirical Validations of Deutsch's Theory of Cooperation and Competition

A number of hypotheses directly follow from Deutsch's theory of cooperation and competition, which may be classified into five major propositions (Deutsch 1985, Deutsch et al, 1988) as follows:

1. A cooperative group exhibits more effective intermember communication compared with a competitive group.

2. A cooperative group demonstrates more friendliness, more helpfulness, and less obstructiveness among its members in its processes.

3. A cooperative group generates more coordination of efforts, more division of labor, more orientation to task achievement, more orderliness in discussion, and higher productivity, if the group task requires effective communication, coordination of efforts, division of labor, or the sharing of resources.

4. A cooperative group produces more agreement and similarity in ideas and more confidence in one's own ideas and in the value that other members attach to those ideas.

5. A cooperative group leads to the defining of conflicting interests as a mutual problem to be solved by collaberative effort, thus facilitating the recognition of the legitimacy of each other's interests and of the necessity of searching for a solution that is responsive to the needs of all.

The empirical validation of Deutsch's theory can be reviewed in two parts: the studies of the dimensions of the characteristics of cooperation and competition and the tests of the effects of cooperation and competition.

The studies of the dimensions

Earlier research in this area can be generally represented by Wish et al's studies (1974, 1976, 1977). These studies took a purely exploratory approach. A multidimensional scaling analysis revealed four dimensions such as "cooperative and friendliness vs. competitive and hostile," "equal vs. unequal," "intense vs. superficial," and "socioemotional and informal vs. task-oriented and formal". The validity of these extracted dimensions was supported by studies of Deutsch (1949a, 1949b), Triandis (1972), Kelley (1979) and Marwell and Hage (1970).

More recent research was largely conducted in the school or business organizational setting (Tjosvold, 1984, 1990). Zhang (1990) used an approach similar to a second order factor analysis. Seven paired categories derived from earlier studies were organized into a rating scheme of classroom climate*. A multidimensional program extracted three dimensions: noisy vs. quiet, egalitarian and informal vs. non-egalitarian and formal, and social-oriented vs. task-oriented. These dimensions seemed matching with those from Wish et al's studies quite well.

Seven equations were formed to regress teacher-self-reported proportions of time used with different teaching modes on the three dimensions. These teaching modes are "total class activity," "small group as part of a systematic program of cooperative learning," "small group not as part of a systematic program of cooperative learning," "lecture," "individualized work," "student discussion," and "other methods." The results indicated that the three dimensions were only sensitive to the proportion of time with which a systematic cooperative learning method was employed. Controlling for multiple equations, a studywise significance level is .038 based on a binomial distribution of F. In other words, if the null hypothesis is true, the probability that one or more equations reach the obtained significance level for that given equation (.0055) is .038. These results strongly suggest that the dimensions have captured the unique characteristics of cooperative learning and the discriminant validity of those dimensions with respect to their potency of measuring the degree of cooperation, which in turn provides further support for Deutsch's theory.

The empirical studies have demonstrated that Deutsch's characterization of cooperation and competition has been consistently supported. Although the basic dimensions obtained vary slightly in different studies, the differences are generally attributable to the idiosyncrasies of each study such as research questions, labelling procedures, levels of classification, research populations, etc. It is to be noted that the above five propositions are not necessarily corresponding to each of the basic dimensions, since they were derived from the generalization of many studies. The application of them, therefore, ought to be situation-specific, if one wants to construct research instruments based on them.

The tests of the effects

There are numerous studies, published or unpublished, in this area after Deutsch's initial research (Deutsch, 1949b; Back, 1951; Berkowitz, 1957; Grossack, 1954; Raven and Eachus, 1963; Johnson and Johnson, 1972; Zajonc and Marin, 1967). These studies can be examined roughly according to such topics as communication, affective processes, group building, and work productivity as consequences of cooperation or competition. For instance, Grossack (1954) found that cooperative subjects showed significantly more cohesive behavior, and also received significantly more communications, more instrumental communications (opinion and information), and less consummatory communications (tension and antagonism). In their study of the effects of communication on game partner's behavior, Pareek and Narendra (1977) demonstrated that communication increased the tendency of a subject to change from competitive moves to cooperative moves in a "maximizing difference" game.

A number of studies found an interesting phenomenon: low-performing individuals are better liked in cooperative than in competitive learning situations (Johnson and

Johnson, 1989). According to Deutsch's theory, in a promotively interdependent relationship, a bungling action would be cathected negatively, that is, a bungling action induces negative feeling or disliking in the partner(s). However, Johnson, Johnson and Scott (1978) in one of their studies found that after individuals worked together cooperatively for ten weeks, and were given a choice of future partners, low-achieving peers were chosen just as frequently as high-achieving peers.

Johnson and Johnson (1989) proposed that there were at least four explanations for this phenomenon. The first was that it was the expectation that the others would act to facilitate the achievement of one's goals that influences liking, not the actual facilitation. Thus whether the low-achieving peers are liked does not depend on whether they have actually facilitated one's goal achievement. The second explanation was that it was the intention on the bungler's part that affected one's liking or disliking rather than the bungler's ability. Tjosvold et al. (1981) found in their study, when future contact in work or social settings was assumed, high-effort partners in their previous collaberative work were better liked than the low-effort partners regardless of their ability. A third possible explanation was that within cooperative situations, individuals were perceived in multidimensional ways. And level of performance was only one of many factors taken into account in forming a positive or negative opinion of the person. The other likely explanation indicated by Johnson and Johnson was that cooperative work produced personal commitment to each other, especially when help and assistance was given to the low-performing individuals. Johnson et al. (1983) tested this hypothesis, and they found that low-achieving minority peers were better liked in cooperative than in individual learning settings.

From this brief review of testing of the affective aspects of cooperation and competition, two theoretical questions arise which in fact relates to theory construction and theory testing in general: (1) to what extent a theoretical model should accommodate relevant intervening variables such that it may claim its validity, predictability or applicability; (2) does a discovery of new moderating or mediating variables imply the disconfirmation of the theoretical model? These concerns were particularly reflected in the controversy between Johnson et al. (1982) and Cotton and Cook (1982) and McGlynn (1982) regarding Johnson et al's meta-analysis (1981).

Johnson et al. (1981) conducted a meta-analysis and reviewed 122 studies on the effects of cooperative, competitive and individualistic goal structures on achievement. These studies had yielded 286 findings. Johnson et al. found that a) cooperation is considerably more effective than interpersonal competition and individualistic efforts; b) cooperation with intergroup competition is also superior to interpersonal competition and individualistic efforts; and c) there is no significant difference between interpersonal competitive and individualistic efforts. In addition, a number of potentially mediating variables for these results were identified.

The arguments were presented around whether it was appropriate to claim the superiority of cooperation over the other two goal structures when potentially there were mediating variables. Cotton and Cook (1982) contended that Johnson et al. had ignored the interaction effects, with mediating variables such as task interdependence, type of task, and type of reward, and thus based their conclusions solely on an undue generalization of the main effects. Johnson et al. replied (1982) that the existing evidence from the 122 studies was simply not sufficiently strong to draw conclusions about the effects of the mediating variables.

So far as we are concerned here, however, the real methodological issue of theory testing revealed in this case is the implicit assumptions embedded in the arguments by Cotton and Cook. They (1982) cited Nisbett and Ross (1980) and warned that there was a perseverance effect with theories of psychology such that overwhelming evidence was required to overturn conclusions once they had been generally accepted. They continued with the remark that "it took nearly 20 years after Deutsch's (1949) research on cooperation and competition to question his findings.

...After this slow progress, let us not be too hasty to make sweeping generalizations again."

If we understand Cotton and Cook correctly, the arguments they have presented imply two methodological assumptions: 1) a theoretical model should be able to predict the future precisely by including all the relevant conditioning variables, or at least the most important ones; 2) the conclusions or hypotheses from the model should only survive when such predictions are confirmed. McGlynn (1982) has pushed those arguments a step further. In his commentary article (1982), McGlynn indicated that "(i)t does no good to know that cooperation is superior summing across all circumstances. What is practical, to paraphrase Lewin, is to have a good theoretical statement that allows one to deduce whether the given circumstances favor the use of one goal structure or another or whether the circumstances can be altered to enhance the effects of a given goal structure." In other words, a good theory is expected to be like a cooking recipe such that all the ingredients and the interactive relationships among those ingredients are carefully specified for ready use.

We would have no hesitation in accepting these arguments if only the reality could be somehow arranged in such a nicely consistent manner that a theory could give unequivocal instructions for practitioners' use in any concrete circumstances. Unfortunately, empirical research has shown tremendous inconsistency in terms of which variables have significant interactive effects with the goal structures. For instance, in addition to the moderating variables Johnson et al. (1981) and Cotton and Cook (1982) have mentioned, research has suggested that "individual accountability" may also have intervening effects in the context of coopeartive learning. However, as Deutsch (1985) has pointed out, on theoretical grounds as well as on the basis of research evidence, there is reason to believe that one of the consequences of successful cooperation is the induction of a heightened sense of responsibility toward one's fellow group members. Thus the repeated experience of successful cooperation is expected to reduce the need for explicit individual accountability. Deutsch (1985) has indicated that "considerable research is needed to identify the conditions that foster the sense of responsibility to others and deter free riding". In other words, while the effects of the first order interaction variable "individual accountability" remains to be examined, the search has been called for of the second order interaction variables--"conditions for sense of responsibility". Thus how should we go on with theory testing while many potential uncertainties regarding moderating variables exist? It is this issue of fundamental methodological importance that we intend to explore next.

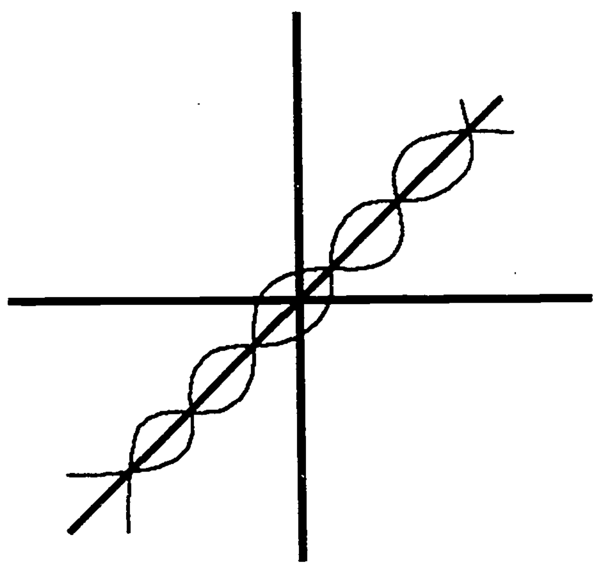
Chapter Two

METHODOLOGICAL OVERVIEW

A Methodological Issue--Levels of theory testing

There is no doubt that theory-data consistency is a basic criterion for theory testing. If a theory consistently makes hypotheses which are inconsistent with data, the theory will have no grounds for being maintained. However, the basis for claiming that the theory is consistent or inconsistent with the data is the first issue we have to address before we test the theory. The history of science has shown that theory testing is an enterprise of endless approximation. In statistical jargon, theory testing is a matter of curve-fitting in both natural and social sciences. Say, an intersection of a theoretical curve and a data point is considered to be a piece of evidence that the theory is consistent with that data point. Given a sample of data points in the continuous case (see figure 1) (Sober, 1988), even if the curve fits every point in the sample space, we can still draw many other curves, in fact an infinite number of curves, which fit the data equally well. Because there are an infinite number of locations that the points in between any two sampled points might potentially reside. This is a well known phenomenon in both social and natural sciences: indeterminacy of fitting functions. Furthermore, if we would like to play a little hair-splitting game, the sampled points, which mathematically are conceptualized as instantaneous points, are no more than statistical averages of small intervals around them measured in an actual research setting. Thus the sampled points themselves are just rough indexes, given they have been accurately measured.

Figure 1.



(In fact we can draw an infinite number of curves passing through the data points.)

However, granted that our theoretical curves equally fit the data, can we then claim all these theories are equally valid and powerful? If we conceive the same issue in a different way, what if the curve-data fitting varies for these theories? Apparently we need further confinements in order to answer the questions properly. In the case of Figure 1, for instance, if we add more dimensions to the graphs, the curve-data fitting pattern is very likely to change. So the well-fitted curves in the two dimensional space might be fitted in a quite different pattern in a three dimensional space, if an effective moderating variable is found. It is also possible that the three dimensional graphs have a misleading fit due to the potential effects from the second order factors, as in the case of the effects of "individual accountability" theoretically conceived above. Or it is still possible that the three dimensional space does not have a decent fit at all due to the misspecification of the first order variables. We do not intend to exaust all the "possibilities" here but rather to borrow an example from Howell (1982) to illustrate and expand the situation a little further.

Given we have a set of raw data:

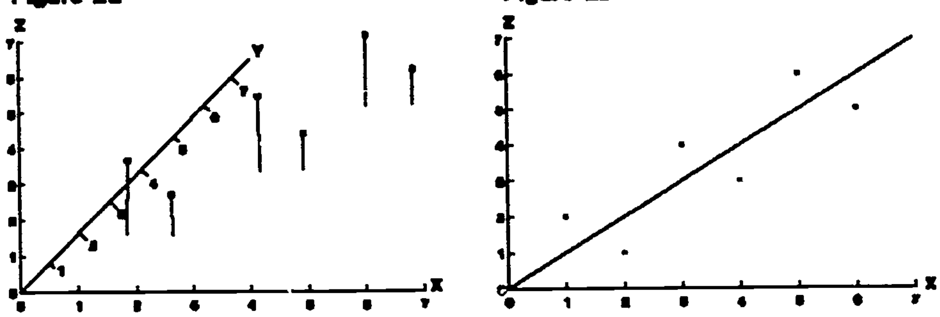
Z	2	1	4	3	6	5
X	1	2	3	4	5	6
Y	2	2	4	4	6	6

Figure 2a represents the three dimensional projection of Z on X and Y. It is clear that for any single value of Y, the slope of the regression line of Z on X is decidedly negative (in fact they are -1). For example, holding Y at 2, then when X equals 1, Z equals 2; and when X equals 2, Z equals

1. The slope is -1. Similarly, when Y is held at 4 and 6 respectively, Z and X are also perfectly negatively correlated. Thus the partial regression coefficient of Z on X controlling Y is -1 which is the average of all the three separate slopes at the three levels of Y. But if we look at Figure 2b where Z is plotted against X, ignoring Y, the slope is positive (.829)

Figure 2a

Figure 2b



a) Z as a function of X and Y; b) Z as a function of X only without controlling Y

With a moderate sample size, a slope of .829 is surely significantly different from zero. Thus we may make the drastically erroneous conclusion that our research hypothesis is consistent with the data.

Things have been fairly obvious at this point: when we claim a theory-data consistency, we only refer to it within a certain scope, that is, at that level of testing our theory is confirmed by the data. The same is true for the disconfirmation of the theory. Thus the inclusion or exclusion of a new variable in theory testing in fact changes the testing to a different level. The conclusions from the testing then must also have varying implications to the theory being tested. The results may confirm or disconfirm some part of the theory depending upon which part the test pertains to. This is at least one of the reasons why the perseverance effect exists. Since most of theory testings are conducted at a local level, and their power of confirming or disconfirming is therefore relatively concentrated at that level. The number of confirmations over the total number of tests (<u>holding the level of testing constant</u>--the author added) constitutes the degree of confirmation (Hempel, 1965), which may suggest whether a theory is worth keeping. It is practically impossible to test a theory at the global level such that the theory can be confirmed or disconfirmed entirely.

Hempel (1965) presented a very interesting case to explain confirmation power in theory testing which is known as the paradox of confirmation. Consider the two sentences:

S_1: All A's are B. ("All ravens are black.")

S_2: All non-B's are non-A. ("Whatever is not black is not a raven.")

Let (a), (b), (c), (d) be four objects such that (a) is a raven and black, (b) a raven but not black, (c) not a raven but black, and (d) neither a raven nor black. Logically, the statements S_1 and S_2 are equivalent. Confirmation of S_1 implies the confirmation of S_2, which Hempel(1965) defined as the equivalence condition. It is no doubt that the observation of "a" confirms S_1, "d" certainly confirms S_2, since these are the direct logical consequences. Thus by equivalence condition, the observation of "d" also confirms S_1; i.e. that anything that is not black being observed is not a raven confirms S_1 too. Consequently, any red pencil, green leaf, yellow cow, etc. becomes confirming evidence for the hypothesis that all ravens are black. This reasoning has led to the conclusion that a white shoe observed confirms that all ravens are black. However, a white shoe observed seems to be entirely irrelevant to the hypothesis we are testing. To straighten this seemingly paradoxical case out, Hempel (1965) employed another example.

Suppose we had a hypothesis that "All sodium salts burn yellow" (All A's are B). We now tested this hypothesis by holding a piece of pure ice into a colorless flame and found that it did not turn the flame yellow. This result would confirm the assertion that "Whatever does not burn yellow is no sodium salt (All non-B's are non-A)." By the equivalence condition, it would confirm our hypothesis. Why does this impress us as paradoxical? Imagine that we did not know what we were burning is a piece of ice, and when we observed that the flame did not turn into yellow and checked back and found that it was not sodium salt but a piece of ice, would we consider the result of this experiment as a confirmation of our hypothesis? The answer is most likely to be "yes".

Hempel (1965) thinks that it is the lacking of prior information (Sober (1988) called background information) about what was burning and the possible consequences of the experiment relative to the hypothesis being tested which enhance the power of confirmation of the test. In other words, it is the uncertainty about the experimental consequences that correlates the power of confirmation of a test. By this formula, the power of confirmation is equivalent to the amount of information that the test may provide. Obviously, the more variables are involved in a test, the greater the uncertainty is regarding the results of the test.

In addition, what we are also interested to know in the real context of theory testing is how the false disconfirmation of S_1 happened at the first place in the above experiment. Thus, another heuristic implication from Hempel's example to our theory testing is that the accidental "disconfirmation" of S_1 was in fact caused by observational (measurement) error--the piece of ice was mistakenly treated as sodium salts! If we had not checked back and re-observed the chemical attributes of the ice, the incident of "disconfirmation" would have been recorded as an anomaly to the proposition being tested. It is quite alarming that in the common practice of theory testing, all the operational variables are treated as error free, that is, they are assumed precisely measuring what they purported to measure. This may be quite misleading.

Confirmation power is closely related to the levels of theory testing. For instance, the power of the observations of 10 black ravens at the same geographic area is much smaller than if the observations were made at very different geographic regions in supporting the hypothesis. Thus by taking geographic characteristics into account, the theory testing is brought to a higher level and the uncertainty about the consequences is also increased. On the other hand, if the interaction effect exists, that is, if the hypothesis "All ravens are black" is not universally true, then the selection of the moderating variables such as geographical altitude, climate, environmental pollution, etc. will greatly affect the power of the testing. In general, a higher-level test accompanies greater confirmation power. (It is not symmetric with the power of disconfirmation.)

A significant trade-off in conducting theory testing at a higher level is that if any mediating or moderating level is not included or not properly specified or the measurement error for the operational variable is large, the results of the test can be misleading. In the situation where a higher order variable does not have direct effects over the lower order variables, i.e. the complete mediation model: X-->M -->Y (James and Brett, 1984), the tests of the effects of the higher order variable will be heavily dependent on the degree that the mediating variables are correctly specified and properly measured. For instance, the attribution model of leader behavior by Green and Mitchell (1979) explicates the relationships between a leader and his or her subordinate in the form as: subordinate's performance (X) ---> leader's attributions of the causes of the subordinate's performance (M)---> leader behavior toward the subordinate (Y). Where the correct conceptualization and measurement of the mediating variable, attribution, to a great extent determines the validity of the empirical conclusions regarding X and Y. The situation with the existence of moderating (interactive) variables is even more complicated. Therefore the higher the testing level is, the more likely that misspecifications occur, since more variables are involved in a single test. Thus careful theoretical guidelines are particularly vital to the theory testing with an integrative approach.

Fortunately, methodological and statistical developments (Bentler, 1986; Bentler 1989; Blalock, 1964; Goldberger, 1964; Joreskog, 1973; Joreskog and Sorbom, 1979; Joreskog and Sorbom, 1989) in the past two decades have provided an excellent solution, which may not be optimal but at least very useful, for the above dilemma. The approach of linear structural relations (LISREL) (Joreskog, 1973; Joreskog and Sorbom, 1989) has demonstrated that a theory can be tested at much higher level, compared to that in the past, while maintaining the theoretical coherence and statistical integrity. Furthermore, this approach takes a confirmatory form such that theoretical guidelines can be incorporated into the model building of the test, thus to reduce the risks of misspecifications and capitalizing on the chance.

LISREL VS. Traditional Approach of Theory Testing--Confirmatory VS. Exploratory

The traditional approach of exploratory nature, Fornell (1982) regarded as the first generation of multivariate analysis, may be best represented by principle component analysis, factor analysis, multidimensional scaling, MANOVA, discriminant analysis and loglinear. In conjunction with the developments of computer science, the traditional approach grew rapidly in the past three decades. Statistical methods such as factor analysis which were quite formidable to scientists and researchers and almost unfeasible even in the forties and early fifties became very popular in the seventies. Along with this change, large data sets could be rather easily handled by ready-made statistical computer packages. In addition, statistical sophistication reduced the number of required assumptions to the few basic ones, which enabled much greater access to the up-dated technology.

For instance, in the early days, the normality assumption had to be routinely checked before an ANOVA could be employed (Still and White, 1981). A deviation from the normality would demand that a transformation procedure be conducted or a nonparametric method be used instead of the F test. After computer simulation techniques became available to the application-oriented statisticians, it was found that a number of statistical methods such as ANOVA and regression analysis were fairly robust against the violation of the normality assumption (Norton, 1953; Boneau, 1959; Havlicek and Peterson, 1974). Robustness means that even if there are departures from normality in the population, the test will have approximately the same power and size as it would if there were no such departures. In recent years, advances in statistics have reached the point such that analytical results have demonstrated that the robustness of normal theory methods is not only applicable in the simpler inferential statistics but also in the analysis of linear latent variate models such as LISREL (Anderson and Gerbing, 1988; Browne and Shapiro, 1988). For example, it was found that if common factors have arbitrary distributions but unique factors still are normally distributed in factor analysis and related models, most asymptotic properties of normal theory methods are retained (Amemiya, 1985; Browne 1985, 1987). In other words, when the sample size is sufficiently large, the violation of normality assumptions for the common factors would not impair the inferential power of the tests. In subsequent studies, the restrictions on the unique factors were relaxed too (Browne and Shapiro, 1988; Amemiya and Anderson, 1985; Anderson and Rubin, 1956).

The progress which was very much embraced has in fact brought with it severe drawbacks. As Fornell (1982) pointed out the relaxation of statistical assumptions meant that many methods which were best suited for descriptive research or a priori hypotheses were often replaced by post hoc interpretation, and thus the empirical precedence took over the theoretical relevance. "Virtually all social sciences received a heavy dose of empiricism; data ruled while theory took a back seat." (Fornell, 1982) A fairly common practice of factor analysis or principle component analysis is to throw a large number of operational variables into the pool and instruct the computer to extract a few factors; then these factors are labeled and some scales are thus constructed. By this purely empirical or exploratory approach, items may be clustered due to many possibilities: sample peculiarities, causal relations, common third variables, systematic response errors (response patterns), or simply chance variance. We believe that this approach is mostly used by unsophisticated researchers for their ad hoc purpose, still the cases are not rare. Thus, under the highly advanced technology, the purely exploratory approach can be quite a sneaky business. As long as there is a large number of variables available, the chance for finding a significant result somewhere is ample. Probably, the exchange of inferential theory and a priori hypotheses for fewer statistical assumptions and a process by which data preceded conceptualization was a price paid for using the exploratory methods developed in the past several decades.

However, this is not a price we have to pay, at least the price may not be so high. Whereas exploratory methods required little in terms of theoretical guidelines, the second generation of multivariate methodology (Fornell, 1982) emphasizes the cumulative aspect of theory development by which prior knowledge may be incorporated into empirical analysis. Such prior knowledge can derive from theory, previous empirical findings, or research design. This new methodology is defined as a confirmatory approach.

Some people may argue that the prior knowledge must be ultimately attributable to empirical studies, thus in essence the new approach is no more than a-few-step-further exploratory. Although this classic debate regarding induction and deduction which started with Aristotle (Mulaik, 1987) has not yet been solved so far, one point may be agreed upon: namely, it is this "a-few-step-further" that introduces the new quality--logical thinking--into our statistical analysis, particularly in the circumstances where our theories can be translated into numerical figures to be built into the statistical models and tested directly. Confirmatory factor analysis and confirmatory multidimensional scaling are good examples, as compared to their exploratory counterparts.

In the exploratory factor analysis, for instance, the common factor "g" can not be uniquely determined by the observed variables. All that we know are the correlations of the observed variables with the g factor, and mathematically, it is possible that a number of distinct variables might be found that have the same pattern of correlations with the observed variables as does the g factor. Thus the g factor in the exploratory approach is no more than a mathematical artifact which needs to be conceptualized. It is in this sense that Mulaik (1987) has recommended that the exploratory approach be taken merely as a hypothesis-generating method.

In the confirmatory approach, on the other hand, certain parameters can be fixed in advance according to prior knowledge, thus building human logical thinking into the statistical procedure. For instance, when fitting a function to a set of data points as we discussed before, we may commit ourselves to certain assumptions that lead us to pick one of the infinitely many generalizing functions based upon prior knowledge derived from previous studies. A test of the fitness of the chosen function will then become the test of the validity of the conceptualization of the previous research. From the probabilistic point of view, a basic advantage that the Bayesian approach posesses over other statistical methods is that it can incorporate prior knowledge to guide the present study. By using a confirmatory approach, causation, as a product of human conceptualization of the relationshiops among variables, can be brought under direct scrutiny by fixing some directions in advance, even when the data are collected in a non-experimental setting.

Of course all the traditional methods are not necessarily purely exploratory. For instance, we may build our prior knowledge into a regression analysis and specify the dependent and independent variables according to our causal assumptions. However, regression analysis can only handle the situation where multiple variables are restricted on one side of an equation. As Fornell (1982) pointed out when multivariance is restricted to either predictor or criterion variables, we do not have a truly multivariate analysis. The common practice of estimating separate multiple regressions for different criterion variables assumes that the correlation matrix of y-variables is an identity matrix, that is, the y-variables are not correlated. This is apparently an unrealistic assumption. Thus regression analysis is inadequate in performing theory testing of high level (integrative level) where multiple dependent and independent variables and moderating/mediating variables have to be analyzed together. Compared to the traditional exploratory approach, the confirmatory approach represented by LISREL has much greater advantage in theory testing.

Major Advantages of LISREL in Theory Testing

From the above methodological discussions, it has become clear that theory testing of greater power requires that the test be conducted at a higher level; a test of higher level requires better theoretical guidelines. Furthermore, psychological constructs are not directly measurable, thus measurement error in the operational variables must be taken into account. To meet all these challenges, LISREL is generally considered to be superior to the traditional methods.

LISREL, also known as the JKW model by Joreskog (1973), Keesling (1972), and Wiley (1973), is a generalization of path analysis from sociology, simultaneous equations from economics, factor analysis from psychology, and the concept of true score in psychometrics (Wright, 1934; Blalock, 1964, 1971; Goldberger, 1964, 1970; Duncan et al, 1968; Bentler, 1976, 1986; Spearman, 1904; Guttman, 1952; Durbin, 1954; Madansky, 1959; Tucker, 1955; Joreskog, 1969; Lawley, 1967). Fornell (1982) edited his two-volume book titled "A Second Generation of Multivariate Analysis" in which LISREL is viewed as the leading trend. The main advantages of LISREL approach in theory testing can be summarized in three aspects:

1. Causal modeling at higher levels

Although statistical methods capable of processing multiple criterion variables and multiple predictor variables were available three decades ago in economics and sociology such as canonical correlations, simultaneous equations and path analysis (Wright, 1934; Simon, 1954; Tukey, 1954; Blalock, 1964; Duncan, 1966), econometric models were rarely applied outside economics and path analysis and canonical correlations were only occasionally utilized in psychology. It was Karl Joreskog (1973) who integrated the econometric, sociometric, and psychometric models and theories and first realized the computer implementation of LISREL (Bentler, 1986; Bollen, 1989).

In addition to processing psychometric factor analysis and econometric simultaneous equations together, LISREL handles both manifest and latent variable structures. By incorporating all the important variables into a single model, a well defined structure of linear relations defined by strong theoretical guidelines can be tested in a confirmatory fashion.

2. Incorporating abstract and unobservable constructs

The typical procedure of causal modeling by LISREL involves first obtaining the latent variables from manifest variables and then estimating the relations between the latent variables. It is this feature of estimating causal relations at the latent level that represents major progress in the methodology of theory testing.

As we have indicated before, manifest variables only rarely correspond in a one-to-one fashion with the constructs of interest to the researcher, thus measurement errors are almost certainly unavoidable. As a consequence, conclusions made based on manifest variables cannot be relied upon, since various theoretical effects will of necessity be estimated in a biased manner and estimated parameters will be very likely to fluctuate from study to study. That is probably one of the reasons that results in social sciences from one study are difficult to replicate by another. Therefore, the main achievements of LISREL in theory testing lie in its ability to separate error from meaningful effects as well as the associated parametric invariance obtainable under various circumstances (Bentler, 1980).

3. Combine and confront a priori knowledge and hypotheses with empirical data

Fornell (1982) stated that if growth of knowledge is a measure of scientific progress, a priori notions must play a dual role in empirical inquiry: as assumptions that are given and as hypotheses that are to be tested. The former role combines theory and data; the latter confronts theory with data.

As a confirmatory approach, the application of LISREL to theory testing utilizes as much a priori knowledge as available to the researcher. Thus the theoretical framework helps identify pertinent variables, permits sources of extraneous variance to be identified and controlled, makes possible careful definition and operationalization of independent variables, and alerts the researcher to potentially important or intrusive interactions (Lipsey and

Pollard, 1989; Chen and Rossi, 1983, 1987). And in particular, LISREL implements those theoretical guidelines in a numerical fashion. As a result, LISREL goes beyond curve-fitting and indicates statistical inference theory can be incorporated in the estimation as well as applied to the theory itself which is embodied in the parametrization of the model being tested.

Chapter Three

RESEARCH HYPOTHESES

Further Developments of Deutsch's Theory

In his Principia, Newton (1953) wrote, "(i)n experimental philosophy we are to look upon propositions inferred by general induction from phenomena as accurately or very nearly true, notwithstanding any contrary hypothesis that may be imagined, till such time as other phenomena occur by which they may either be made more accurate or liable to exceptions." In other words, theory testing provides the impetus for theory development in empirical sciences.

Some application oriented modern social scientists further suggested that theory testing be conducted in applied settings (Leventhal, 1980; Rodin, 1977; Collins et al. 1980). As Sanford (1970) noted, for Lewin, involvement with practical problems was a never-failing source of theoretical ideas and knowledge of fundamental social psychological relationships. Research that attempts to provide answers to questions posed by the real world can lead to the formulation of general principles, in other words, theory (Rodin, 1985).

Within the traditional framework of theory testing, Deutsch's theory of cooperation and competition has been modified and expanded into a theory of cooperation and conflict resolution. A number of relevant new variables were identified; predictive propositions regarding more general social psychological aspects have been formulated (Deutsch, 1987; Johnson and Johnson, 1987, 1984; Sarason et al. 1983; Seta, Paulus and Schkade, 1976; Sharan, 1980; Slavin, 1984; Slavin and Karweit, 1981; Tjsovold, 1982; Tjsovold, Johnson and Johnson, 1984). It is proper to say that the accumulated knowledge in the past four decades has well prepared us to conduct an integrative test of Deutsch's theory.

The focus of the present study

The objective of this study is conceived as two-fold: the evaluation of the intervention of cooperative learning and conflict resolution in an inner city high school and the testing of Deutsch's theory of cooperation and conflict resolution. As we stated previously, Deutsch's crude law emphasizes the reciprocal relations between institutional variables and outcome variables such that the changes on the former would bring about and in turn be reinforced by the changes on the latter. Thus the focus of the study lies in the impact of institutional changes induced by the intervention.

One of the logical outcomes of the intervention of cooperative learning and conflict resolution is the change of social climate. By the term social climate we refer to the interpersonal aspects of organizational environment under which individuals' behavior is affected. Rodin (1985) has pointed out that the study of applied social problems has benefited greatly from the orientation of social psychology toward the social environment. Understanding that situational settings and the behaviors that occur there are mandated by the norms and rules of social process rather than by individuals' personalities provides an important focus for intervention and for social changes (Rodin, 1985).

A number of studies have demonstrated that situational factors tends to promote certain types of behavior which in turn reinforce the social norms and rules (Langer and Rodin, 1976; Rodin and Langer, 1977; Schulz, 1976). For instance, in their study of nursing homes, Wack and Rodin (1978) considered that situational factors forced nursing homes to promote dependence and passive behavior. Rules promulgated to raise the standards of nursing care and to regulate the administration of nursing homes supported by Medicare and Medicaid created environments that have had the unintended consequences of encouraging dependency and helplessness in patients, making work in nursing homes unrewarding for staff, and discouraging residents from improving enough to lead independent lives again. This dependency, created and supported by contextual variables, in turn leads to further debilitation and ill health. In other words, norms and rules, introduced by social intervention, change the social climate which causes either positive or negative social psychological consequences and reinforces the positive or negative cause-effect cycle.

Recognizing the potential impact of environmental factors, a theory based intervention can be employed to lead social psychological changes in the desired direction by establishing a set of new norms and rules in the target institution. Therefore, the main goal of the intervention of cooperative learning and conflict resolution is to introduce norms and rules which are expected to facilitate positive changes in the social climate of a school system.

As we stated before, the intervention of cooperative learning and conflict resolution basically can be viewed having two components: constructive conflict resolution and effective group learning. Deutsch (1987) articulated the key rules--Deutsch's ten commendments--which can be utilized to generate a systematic training program for an intervention of conflict resolution, and in fact, they have been used as curriculum guidelines for the training in the present intervention, which we quote as follows:

1. Knowing what type of conflict you are involved in.

2. Respect yourself and your interests, respect the other and his/her interests.

3. Distinguish clearly between "interests" and "positions".

4. Explore your interests and the other's interests to identify the common and compatible interests that you both share.

5. Define the conflicting interests between oneself and the other as mutual problems to be solved cooperatively.

6. In communicating with the other, listen attentively and speak so as to be understood: this requires the active attempt to take the perspective of the other and to check continually one's success in doing so.

7. Be alert to the natural tendencies to bias, misperceptions, misjudgments, and stereotyped thinking that occur in oneself as well as the other during heated conflict.

8. Develop skills for dealing with difficult conflicts so that one is not helpless nor hopeless when confronting those who are more powerful, those who don't want to engage in constructive conflict resolution, or those who use dirty tricks.

9. Know oneself and how one typically responds in different sorts of conflict situations.

10. Finally, throughout conflict, one should remain a moral person--i.e., a person who is caring and just--and should consider the other as a member of one's moral community--i.e., as someone who is entitled to care and justice.

The ten commendments have been rephrased into a set of teaching topics for easy learning and for the purpose of evaluation of the intervention, i.e. being employed as the manipulation checks. The topics are: active listening, "I" messages, needs versus positions, negotiable versus non-negotiable, individual conflict style, putting yourself in the other people's shoes, anger and violence, reframing the issues in conflicts, criticize ideas and not people, and "win-win" solutions to conflicts.

Correspondingly, a set of standards are introduced for the intervention of cooperative learning (effective group learning) as well: sharing ideas together, sharing bonus points, learning how to work together as a group, group work evaluation, dividing the lesson so that each student had to learn a piece of it and to teach it to the fellow students, differntial roles in the group, and agreement on answers within the group.

The whole essence of the intervention of cooperative learning and conflict resolution is that through systematically introducing the new rules and norms, the promotive goal interdependence and the promotive means interdependence (Johnson and Johnson, 1989) may be established so that positive changes in school social climate and students' social psychological well being will follow. Testing this fundamental prediction will suffice as our research goals.

Conceptualized relations of the relevant variables

1. Indicators of school social climate

It is almost natural to consider social support as the key variable indicating the school social climate, since the whole program of cooperative learning and conflict resolution aims at building mutual understanding, mutual trust and mutual help (Deutsch, 1973, 1985, 1987) among school personnel and students. While controversies still exist regarding the functions of the moderating variables in the meta-analysis by Johnson and Johnson (1981), the relationships between promotive interdependence and social support has been consistently identified. As Johnson and Johnson (1989) indicated "a large number of experimental studies relating social interdependence and social support provides strong evidence that allows causal inferences. The high quality studies especially give this body of work a validity that is unusual in the study of social support. The diversity of subjects, settings, age levels, and operationalizations of social interdependence and social support gives this work wide generalizability."

Social support has been defined as having access to significant others who can provide information indicating love, esteem, and mutual obligation (Cobb, 1976; Kaplan et al., 1979). According to Silver and Wortman (1980), studies suggest that social support has several major components, including the expression of positive affect; the expression of agreement with or acknowledgment of the appropriateness of a person's beliefs, interpretations, or feelings; the provision of material aid; and the provision of information that the distressed person belongs to a network of mutual help and obligation.

Johnson and Johnson (1989) classify social support as having four main dimensions: 1. <u>emotional concern</u> such as attachment, reassurance, and a sense of being able to rely on and confide in a person, all of which contribute to the belief that one is loved and cared for; 2. instrumental aid such as direct aid, goods, or services; 3. information aid such as facts or advice that may help to solve a problem; 4. appraisal aid such as feedback about degree to which certain behavioral standards are met (information relevant to self-evaluation). According to the Johnsons (1990), a social support system consists of significant others who collaboratively share people's tasks and goals and who provide individuals with resources that enhance their well-being and/or help them mobilize their resources in order to deal with the particular stressful situation to which they are exposed.

In addition to social support, school crime and victimization may be treated as supplementry indicators of school social climate. By employing multiple indicators of social climate, we expect to gain differential evidence for the hypothesized effects of the intervention, if the intervention does yield results in the right direction.

2. Previous findings of the effects of cooperation and conflict resolution on social support

According to the meta-analysis by Johnson and Johnson (1981, 1989), cooperation generally produces greater social support than competitive or individualistic efforts. Seventy-two percent of the findings were statistically significant in the favor of cooperation while only 1 percent favored competition. Fifty-nine percent favored cooperation over individualistic efforts and only 4 percent favored individualistic efforts. When the findings were weighted so that each study counted the same and the sample sizes were equalized, the average cooperator experienced social support at about three-fifths a standard deviation above the average person working in competition and seven-tenths a standard deviation above the average person working individually (Johnson and Johnson, 1989).

The mechanism underlying these effects has been explicated by many researchers (Bonoma, Tedeschi and Helm, 1974; Bridgeman, 1977; Crombag, 1966; French, 1951; Fay, 1970; Grossack, 1954; Jecker and Landy, 1969; Lerner and Simmons, 1966; Johnson and Ahlgren, 1976; Johnson and Johnson, 1983; Johnson, Johnson and Anderson, 1978, 1983; Johnson, Johnson, Buckman and Richards, 1986; Tjosvold and Johnson, 1978;). Deutsch (1949a, 1949b, 1958, 1960, 1962, 1973) has indicated that promotive interdependence facilitates communication, elicits mutual help, induces positive feelings and attitudes, builds mutual trust, enhances mutual commitment and sense of responsibility among cooperators.

Although social support has been a variable which was employed in a number of studies, very few studies have taken it as the outcome variable. Thus the assumed mechanism underlying the relationship between social support and cooperation/conflict resolution remains to be investigated. Johnson and Johnson (1989) have advocated their tentative interpretations labled as social judgment theory. Unfortunately, the theory is lacking of necessary coherence, and needs empirical support and further elaboration. However, the lacking of the understanding of the functions regarding potential mediating processes does not have to impede our present study. Since as we described above, the positive association has been consistently demonstrated and the existence of the causation--either direct or indirect--has been strongly suggested.

3. Indicators of outcome variables and the relationships among them

Probably if not the most, at least among the most relevant and important outcome variables of social psychology for the contemporary American adolescents are mental health, physical health and achievement. Two more variables may be treated as both outcome and instrumental, that is, self-esteem and locus of control.

Mental and physical health There is a rapidly increasing number of studies regarding mental and physical health in the last 16 years in social, clinical and counseling psychology. The PsychLit database shows that by the key word search, there are 13,445 articles involved with "health" for the period of 1974-1982. Among them 7,881 were related to "mental health" and 211 to "physical health". But for the period of 1983-1991, the number of articles about "health" increased up to 31,079, and among them 12,625 were "mental health" related and 679 "physical health" related. The number of studies relating social support to mental and physical health increased from 62 to 586. These figures to some extent suggest the trend of research interests in this area.

According to the definition provided by the World Health Organization (1948), adopted by Rand Cooperation in their Health Insurance Study (HIS) (Ware, et. al. 1979, 1980), health is a state of complete physical, mental, and social well-being and not merely the absence of disease or infirmity. Three components were identified for the measurement of health: physical, mental, and social health. Contemporary theory and definitions of the different components indicated that each of the three major components are comprised of distinguishable constructs that could be measured separately (Ware et al., 1980).

By convention (Bradburn, 1968; Brook et al., 1980; Ware et al., 1979, 1980; Veit and Ware, 1983), physical health may be defined in terms of functional status, namely, the performance of or capacity to perform six categories of activities (self-care, mobility, physical, role, house hold, and leisure) that are normal for individuals in good physical health. Mental health refers to the psychological manifestations of affective and anxiety disorders, positive well-being, and self-control in contrast with somatic or physiological manifestations of these states. Social health can be viewed in terms of interpersonal interactions and activities indicating social participation.

Although physical, mental, and social health are conceptualized as different aspects of general health, they are both theoretically and operationally interrelated. Changes in mental health status may cause or reflect physical illness and may often related to changes in social well-being. Environmental events or stress often bring about changes in physical, mental, or social health or some combination of these components. It is also reasonable to assume that changes in any aspect of health should be reflected in personal ratings of health in general, which are expected to capture the factor of general health status common to all components.

As the growth of concerns of quality of life, measurement of mental health in general populations has expanded from an almost exclusive focus on negative states or symptoms (e.p., sadness, tension) to include positive states (e.p., happiness, interest in life) (Ware et al., 1979). Due to the relative scarcity of research in the area of social health and its relationship with social support, cooperation and conflict resolution, the present study uses only mental and physical health as the indicators of health outcomes.

Self-evaluation By self-evaluation we refer to the two most commonly used variables: self-esteem and locus of control. **Self-esteem** is defined (Coopersmith, 1967) as the evaluation a person makes and customarily maintains with regard to him- or her-self. The importance of self-esteem has been documented throughout almost all the fields of social sciences. Compared to the findings of the positive effects of self-esteem, the negative influence of poor self-esteem or loss of self-esteem received even broader attention (Wortman and Dunkel-schetter, 1979; Suls, 1982; Block and Thomas, 1955; Coopersimth, 1967; Rosenberg, 1965).

Research has found that low self-esteem is related to emotional problems. Individuals with low self-esteem are more likely to report that they are troubled by anxiety, nervousness, insomnia, unhappiness, and psychosomatic symptoms (Block and Thomas, 1955; Coopersmith, 1967; Fitts, 1972; Howard and Kubis, 1964; Rosenberg, 1965). Low self-esteem is also related to poor academic achievement. O'Malley and Bachman (1979) found a significant correlation between low self-esteem and low grades in school. This phenomenon may be due to the fact that people with lower self-esteem set lower goals for themselves, lack confidence, and assume that they will not succeed even if they tried (Coopersmith, 1967; Purkey, 1970).

A closely related concept to self-esteem is the concept of perceived control. It is beyond the scope of the present study to depict the relationship between self-esteem and locus of control. Nonetheless, it is reasonable to assume that the degree of self-esteem leads to a corresponding level of perceived control, that is, a person who values him- or her-self highly in a given area tends to perceive more internal control in that area. In other words, whether or not a person believes that she or he can make a change by her or his own action depends upon the self-judgment and the evaluation of significant others about her or his ability. It is in this sense that control and predictability of changes or outcomes typically occur together (Geer et al., 1970; Glass et al., 1973).

The relationship between loss of control and the onset of poor health has been the focus of substantial research in recent years (Rodin, 1985). Stimulated by the early, intriguing work of Schmale and Iker (1966), many studies have pursued the question of how feelings of loss of control might be related to negative health outcomes and indeed might stand as a central mechanism in the etiology of disease. In their study, feelings of helplessness and hopelessness among the patients who had suspicious cancerlike symptoms are associated with the actual chance of developing cancer.

Investigators have considered how control mediates the effects of major life changes on health. Glass (1977) compared hospitalized patients with nonhospitalized people in order to assess whether individuals who experience a series of uncontrollable, undesirable life events exhibit greater illness. He found that hospitalized patients reported more uncontrollable losses during the the one-year period prior to hospitalization. This finding, however, was based upon retrospective data, thus suggests only the association between the feelings of loss of internal control and physical well-being. Using both retrospective and prospective designs, Mullen and Suls (1982) found that uncontrollable events were related both to future psychological and physiological distress and symptomatology. Husaini and von Frank (1985) indicated that locus of control as a measure of personal coping resources had the most marked effect on depression, it proved to be as good a predictor of depression as was prior depression.

A number of studies have also demonstrated that the relationships between locus of control and mental and physical health are reciprocal (Brody, 1980; Bandura et al., 1980; Rodin and Janis, 1979). It is certainly plausible to assume that when people are confronted with mental or physical difficulties, their feelings of control over the world are also threatened.

Achievement The third class of the outcome variables of the present study is productivity, i.e. students' achievement. If it is correct to assume that efforts and persistence are at least as important for predicting and controlling achievement- related behavior as are inferences of ability (Darley and Goethals, 1980; Deaux, 1976; Feather and Simon, 1973; Valle and Frieze, 1976), students with internal locus of control should have better school performance given other variables are held constant.

Effort or persistence is normally perceived as an internal cause of achievement. Among the earlier studies, Coleman et al. (1966) has documented that the achievement of nonwhite students was best predicted by a measure of a child's belief that academic outcomes were determined by his or her own efforts.

Corresponding to the Coleman report, a number of studies have related school achievement to locus of control. Franklin (1963) found internality related to the amount of time that high school students spend in doing homework. James (1965) reported that internals were more persistent in their attempts to solve complex logical puzzles. In a study using large samples of eighth- and eleventh-graders from the Chicago schools, Lessing (1969) found that the Sense of Personal Control, as assessed by Strodtbeck's Personal Control Scale (Strodtbeck, 1958), was correlated with grade point averages even when IQ scores had been statistically partialled out; that is, among both the more and less intelligent pupils, sense of personal control was related to actual school achievement. Harrison (1968) likewise found that a sense of personal control allowed for some prediction of success in school regardless of the socioeconomic status of chilren's homes.

Later research has turned to investigating moderating variables such as sex and classroom structure (Dweck, 1975; Dweck and Bush, 1976; Dweck et al., 1978; Wright and DuCette, 1976; Arlin, 1975; Harpin, 1980). Although the existing findings on the relationship between achievement and locus of control are not very consistent (Katz, 1967; Nowicki and Roundtree, 1971; Stephens, 1973; Little and

Kendall, 1978), Weiner (1979) has indicated that the choice to engage in achievement activity is mediated by internal-variable factors, such as effort, that generate positive feelings; second, persistence despite failure is said to be more likely if the causes of failure are seen as variable.

The relationship between achievement and self-esteem has been studied in three ways: (1) on the basis of the principle of reflected appraisals (Rosenberg, 1979), self-esteem may be predicted from academic achievement and evaluations of significant others. Scheirer and Kraut (1979) argued that self-concept change is likely to be an outcome of increased achievement rather than a necessary variable for achievement to occur; (2) on the basis of self-consistency theory (Jones, 1973), one may predict that students with low academic self-concept will avoid situations that could alter their self-concept, and hence make less effort to do well in school. Thus, achievement in turn may be affected by self-esteem; (3) Marsh and Parker (1984) proposed a dynamic equilibrium model suggesting that academic achievement, self-concept , and self-attributions are interwoven in a network of reciprocal relations such that a change in one would produce changes in the others to reestablish the equilibrium; (4) Maruyama et al. (1981) indicated there may be "third variables" (e.g. ability) that are responsible for both self-esteem and achievement.

However some recent studies (Skaalvik and Hagtvet, 1990; Hoge et al., 1990) utilizing a causal modeling approach have suggested that the causal direction is from achievement to self-esteem. Skaalvik and Hagtvet (1990) found in their longitudinal study of two cohort groups that the notion of a predominant direction of causality from academic achievement to global self-esteem via self-concept of ability was supported. Hoge et al. (1990) reported that school climate and evaluations by teachers had significant effects on self-esteem. And grades were more important for discipline-specific self-esteem than for global or academic self-esteem.

4. The relationships of social support and the outcome variables

Social support and health

Considerable studies have been conducted investigating the impact of social support on mental and physical health (Cassel, 1976; Cobb, 1976; Cobb and Kasl, 1977; Dean and Lin, 1977; Kaplan et al., 1977; Suls, 1982; Lin et al., 1979; Silver and Wortman, 1980; Kahn and Antonucci, 1981; Brickman et al., 1982; Lieberman and Mullan, 1978). However, evidence of both positive and negative effects are largely available (Rodin, 1985; Wallston et al., 1983; Suls, 1982).

Despite the possibility of negative effects of social support, however, there is strong consensus that, in general, support systems have both preventive and ameliorative effects, especially in the area of health (Rodin, 1985). There are many studies suggesting that people who are part of social networks are less likely to be negatively affected by stressful life events and less likely to become ill. Moreover, there is some evidence that support systems facilitate coping and recovery if people do become ill (Suls, 1982; Kessler, 1982; Wallston et al., 1983).

One of the intensively studied topics is the stress-buffering effects of social support (La Rocco et al., 1980; Norbeck and Tilden, 1983; Williams et al., 1981; Thoits, 1982; Cohen and Wills, 1985; Kessler and McLeod, 1985). In particular, the buffering effects refer to the phenomenon that individuals with a strong social support system are better able to cope with major life changes, while those with little or no social support are relatively vulnerable to these changes (Thoits, 1982). Research has shown that a favorable social network has been linked to fewer complications during pregnency (Nuckolls et al., 1972), faster recovery from illness (Egbert et al., 1964; Bloom et al., 1978; Leavy, 1983), and fewer health problems during stressful periods of life (Cobb, 1976; Gore, 1978). There is also evidence that when levels of stress are low, the enlistment of social support may be unnecessary, while when stress is high social support is a coping mechanism that individuals resort to (Jenkins, 1979; Lin et al., 1979; Seers et al., 1983).

Social support is assumed to have health-promoting effects (Berkman and Syme, 1979; Blazer, 1982; Lynch, 1977;

Mitchell et al., 1982). Virtually all health surveys indicate that married people report being healthier and happier than unmarried people (Comstock and Slome, 1973; Greenley and Mechanic, 1976; Wan, 1976). Lynch (1977) use mortality figures to argure that the unmarried and socially isolated die prematurely because of loneliness. Among college students, social environments characterized as low in community and cohesion, student participation, and social activities resulted in increased health problems (Moos and Van Dort, 1976). Kahn (1981) has indicated that interpersonal relationships are major determinants of how the objective work environment affects an individual's subjective state, psychological responses, and ultimate health and disease.

A number of studies demonstrated that social support has the effects of assisting mental adjustment (Andrews et al., 1978; Bowlby, 1969, 1973, 1980; Cohen and Syme, 1985; Henderson et al., 1980; Sandler, 1980; Wallston et al., 1983). Cobb (1976) noted that social support may protect people in crisis by facilitating coping and adaptation. Bowlby (1980) showed that children with supportive parental figures become self-reliant, learn to function as supports to others, and have decreased likelihood of psychopathology in later life. Vaillant (1974, 1977) found that a supportive early family environment was correlated with positive adult adjustment and lack of psychiatric disorder.

However, counterproductive or ambiguous effects of social support have also been widely documented (Lieberman and Mullan, 1978; Pearlin and Schooler, 1978; Caplan et al., 1976; Caplan et al., 1980; Hartley et al., 1976; Smith and Glass, 1977; Bucke and Parke, 1972; Epley, 1974). Methodological issues have been discussed regarding the inconsistency of the empirical findings (Suls, 1982; Kessler, 1983; Thoits, 1982; Wallston et al., 1983; Vinokur et al., 1987). Identified problems with the present literature are indicated such as (1) conceptualization issues For instance, the theoretical boundry of life events, which are conceptualized as mental or physical stressors, is not specified (Thoits, 1982). (2) measurement issues For instance, Vinokur et al. (1987) suggested that there were two other factors affecting perceived support in addition to interpersonal transactions: personal outlook and transient affective states. People who are poorly adjusted or in ill health may underestimate the amount of support available to them (Suls, 1982). (3) design and analytical issues Results are largely collected from correlational studies, which seems to be unable to isolate causes from effects, especially when a study is conducted retrospectively (which is true in a good proportion of the studies) (Depner and Wethington, 1984; Suls, 1982; Wallston et al., 1983).

Although there exist ambiguity in results and methodological issues in research, the majority of previous empirical findings have indicated the positive effects of social support, strong or modest, on both mental and physical health (Rodin, 1985; Suls, 1982; Thoits, 1982).

Social support and self-evaluation

Since self-esteem is an evaluative psychological apparatus which is affected by the opinions of both oneself and significant others, it would be reasonable to assume that change of self-esteem is contingent upon the status a person posesses in his or her social network. Wortman and Dunkel-Schetter (1979) found self-esteem may be lowered by the impact of ambiguous or negative feedback from other prople. In the situation where people have serious, debilitating illnesses or experiences such as rape, the likelihood of receiving ambiguous or negative feedback from their social network may increase. Thus with time, people may come to internalize the views that they perceive others to hold and begin to feel shame, guilt, self-blame, self-derogation, and even self-hatred.

Fisher and Nadler (1974) suggested that people who receive aid and support may experience it as evidence of failure, inferiority, and dependency, which threatens their self-esteem. Suls (1982) indicates that people who have high self-esteem appear more disturbed by the receipt of assistance than low-self-esteem persons.

It seems that the present literature has documented more studies of negative than positive effects of social support on self-esteem. However, in the situations where social support constitutes a threat to the recipients (Suls, 1982), the content of social support is basically instrumental, thus it may convey the message that the recipients of the kind of support are in a needy position and dependent on the support givers. It is reasonable to assume that social support such as positive informational support, for example, encouragement and praise, may bring about different consequences. In particular, when social support does not deprive the receipients of their self-pride, the above detrimental effects are not likely to occur. Furthermore, it is also conceivable that when the roles of giver and taker are exchangable, namely, the support is mutual to both parties, the negative effects of receiving support should be less prevalent. In fact, it has been documented that self-esteem may be enhanced in the situation of cooperative learning where students engage in the relationship of promotive interdependence (Slavin, 1983; Sharan, 1980; Johnson and Johnson, 1983, 1989). In a cooperative learning classroom, students receive help and support from each other, while everybody is also held accountable for the overall group performance (Deutsch, 1949a, 1949b; Johnson et al., 1986).

A number of researchers have recently reported a positive relationship between social support and self-esteem (Pyskoty et al., 1990; Newcomb, 1990; Mallinckrodt, 1989; Cramer, 1988; Hoffman et al., 1988; DeLongis et al., 1988). Oei and Zwart (1990) found the instrumental function of self-esteem such that effective social support depended on the person helped being able to interpret the aid appropriately so as to feel secure within his or her self-esteem. Similarly, Conn and Christopher (1989) indicated that seeking social support was positively associated with self-esteem, self-efficacy, and optimistic beliefs about the consequences of seeking support. Newcomb (1990) found that early self-esteem increased later support from peers for both sexes and support from adults for girls, which suggested the other direction of the linkage, from self-esteem to social support.

The instrumental functions of locus of control in mobilizing and utilizing social support have been consistently demonstrated (Cummins, 1988; Caldwell et al., 1987; Fusilier et al., 1987; Rivicki and May, 1985; Hibbard, 1985; Lefcourt et al., 1984; Ganellen and Blaney, 1984; Echenrode, 1983). In general, findings suggest that individuals with a strong sense of personal control also possess beneficial support systems in the presence of stressful situations (Revicki and May, 1985). Hibbard (1985) suggested that effective use of social network resources is moderated by locus of control. Having more social ties, being more trustful of others, and perceiving more control facilitate the effective use of social resources. A larger social network is more important for those less able to use the resources available in their netwroks effectively, while among those more able to mobilize support from their networks, size is less important. Similarly Lefcourt et al. (1984) have found that the moderating effect of social supports largely occured among subjects who were less affiliative and more autonomous.

Social support and achievement

Compared to the above described relationships, research literature in the area of social support and achievement is relatively sparse. Achievement is conventionally perceived as determined by personal traits such as cognitive ability and persistence. Whether and to what extent a person can draw upon the external resources available in his or her social network helping him or her achieve deserve further research attention.

Given previous findings, a positive relationship between social support and achievement has been suggested (Haynes and Comer, 1990; Haynes, 1990; Pfannenstiel and Seltzer, 1989; Mallinckrodt, 1988; Cauce, 1986; Crouse 1985; Tracey and Sedlacek, 1985; Blumberg and Flaherty, 1985; Huang et al., 1983; Ruby and Law, 1982). The findings may be classified as follows: (1) material aids from a supportive network have positive impact. Corcoran et al., (1987) studied the effects of support systems (tangible and social) on scholarly productivity. Their results showed that the primary predictive variables related to a composite of scholarly productivity are tangible in nature. No significant relationship was found between "social" (relative to tangible) supports and scholarship. (2)

Parental support is an important dimension contributing to school achievement. Watson et al., (1983) found a significant relationship between parental support and achievement, with neighborhood support having little effect if parental support was absent. Home support was effective regardless of income or educational level of parents. Home support was particularly helpful in the area of expressive language, receptive language, and visual recognition skills measured by Cognitive Skill Assessment Battery. (3) Social support may affect achievement through mediating the negative effects of life events. In their study of children's achievement in single-parent families, Roy and Fugua (1983) found that an adequate social support system may mediate the negative effects that have been reported of single-parent family status on children's academic performance.

Although some studies failed to demonstrate the positive relationship between social support and achievement (Mitchell et al., 1983), in general, social support is perceived to be beneficial to school achievement (Angle et al., 1983). However, the causal directions between social support and achievement are not clear so far.

Research Hypotheses of the Present Study

If Deutsch's theory of cooperation and conflict resolution is valid, the school social climate is then expected to be affected by the intervention, which will in turn have impact on students' mental and physical health, self-esteem and locus of control, and probably, academic achievement. Furthermore, if Deutsch's crude law is valid, the changes in the social psychological aspects of the students will lead to further changes in the school social climate, which constitutes a reciprocal causal loop.

One of the methodological problems associated with previous research is that most data were collected at the same temporal point, and some of the studies even produced data retrospectively, thus causal directions were less justifiable. Since what the outcome variables measured were mostly stable traits, a time lapse should be allowed for the hypothesized effects to occur. In the present study, lagged variables will be incorporated into the models. By means of the intervention of cooperative learning and conflict resolution indicated previously, we will test the following hypotheses to evaluate the results of the intervention and to verify Deutsch's theoretical claims.

Hypothesis 1: The intervention of conflict resolution will increase the student's social support. The magnitude of the effect will depend upon the "dosage" of the intervention received by the student and the base-line level of the measures of the student's social support.

Hypothesis 2: The intervention of cooperative learning will increase the student's social support. The magnitude of the effect will depend upon the "dosage" of the intervention received by the student and the base-line level of the measures of the student's social support.

Hypothesis 3: The intervention of conflict resolution will reduce school crime. The magnitude of the effect will depend upon the "dosage" received and the base-line level of the measures of school crime.

Hypothesis 4: The intervention of cooperative learning will reduce school crime. The magnitude of the effect will depend upon the "dosage" received and the base-line level of the measures of school crime.

Hypothesis 5: An increase in the student's social support will have positive effects on the student's positive mental state and will reduce the students' negative mental state.

Hypothesis 6: An increase in the student's social support will enhance the student's physical health.

Hypothesis 7: An increase in the student's social support will have a positive effect on the student's self-esteem.

Hypothesis 8: There will be reciprocal relations between a student's social support and his or her internal locus of control such that an increase in a student's social support will lead to an increase in the student's internal locus of control and vice versa.

Hypothesis 9: An increase in a student's self-esteem will lead to a decrease in the student's negative mental state and an increase in his or her positive mental state.

Hypothesis 10: An increase in a student's self-esteem will increase the student's internal locus of control.

Hypothesis 11: There will be reciprocal relations between a student's locus of control and his or her mental states such that an increase in his or her internal locus of control will lead to a decrease in his or her negative mental state and vice versa.

Hypothesis 12: Poorer physical health will have a negative effect on internal locus of control.

Hypothesis 13: Higher Internal locus of control will lead to higher academic achievement.

Hypothesis 14: A student's school achievement will positively affect the student's self-esteem.

To test Deutsch's crude law, lagged variables will be used to detect effects or associations of prior social psychological outcomes upon intervention variables, social support, school crime, and a student's victimization. Utilizing lagged variables we also intend to examine the stability or persistence of the social psychological states.

Hypothesis 15: The prior level of a student's internal locus of control will positively affect the amount of social support available to the student or utilized by him or her at a later stage.

Hypothesis 16: The prior level of a student's mental health will be positively associated with the social support received by the student at a later stage.

As we have stated, school crime and victimization are chosen as alternative indicators of school social climate.

We assume that school crime affects individual students through victimization rather than directly impinges upon students' health, self-evaluation, or achievement. Therefore, the following hypotheses will also be tested:

Hypothesis 17: The intervention of conflict resolution will reduce student victimization.

Hypothesis 18: Increased social support will reduce a student's victimization.

Hypothesis 19: School crime will positively affect a student's victimization.

Hypothesis 20: A student's victimization will reduce the student's positive mental state and increase his or her negative mental state.

Hypothesis 21: A student's victimization will have a negative impact on the student's physical health.

Hypothesis 22: A student's victimization will lower the student's self-esteem.

Hypothesis 23: A student's victimization will decrease the student's internal locus of control.

Chapter Four

METHODS

The analytical design of this study has two parts corresponding to our research objectives: 1. causal modeling of the structural relationships among the relevant variables to test the hypotheses and predictions of Deutsch's theory of cooperation and conflict resolution as stipulated above; 2. effect analysis to provide supplementary information for the LISREL models and to evaluate the intervention of cooperative learning and conflict resolution introduced to the inner-city alternative high school.

Design of LISREL Modeling

Deutsch's crude law says that the characteristic processes and effects elicited by a given type of social relationship also tend to elicit that type of social relationship (Deutsch, 1985). To some extent, the present study of testing Deutsch's theory and the intervention project have been designed in a corresponding fashion. The crude law implies a progressive reciprocal loop between the said characteristic processes and their outcomes. What we intended to test in this study were the two consecutive leaps: from the characteristic processes elicited by the interventions, namely, the affected structure of interpersonal relationships within each of the experimental sites which is represented by social support, school crime, and a student's victimization, to the social psychological outcomes; and from the social psychological outcomes to the changing structure of interpersonal relationships. (We occasionally use interpersonal relationship and school social climate interchangably in this thesis.) The second and third waves of data were originally planned to be used for this design which may be depicted as the following diagram:

INTERVENTION Second wave ——> SOCIAL CLIMATE Second wave ——> OUTCOMES Second wave ——>

SOCIAL CLIMATE Third wave ——> OUTCOMES Third wave ——>......

Given this design, the analytical procedure may either focus on the first loop or on the second loop; that is,

SOCIAL CLIMATE Second wave ——> OUTCOMES Second wave ——> SOCIAL CLIMATE Third wave

or

OUTCOMES Second wave ——> SOCIAL CLIMATE Third wave ——> OUTCOMES Third wave

Unfortunately, the interventions ended prematurely due to administrative reasons, and the planned third-wave data collection was not completed. Thus, instead of using the third-wave data, we used the first-wave data, that is, the data collected prior to the interventions. In other words, we will use the potential pre-existing relations between self-evaluation, mental health, and physical health and the social climate variables and/or intervention variables to roughly assess the effects that the outcome variables might have on the intervention and/or social climate. However, by using the first-wave data to simulate the second leap in the above chart, we will underestimate the relationships between the outcome variables and the intervention variables or climate variables, since the co-variation is expected to be less systematic than otherwise might be if the post- and post-post-data had been used. Furthermore, any relationship estimated may only indicate an association rather than a causal linkage. It should be noted that the estimation of the effects of the intervention variables on the intervening and outcome variables, namely, the first leap, is the main focus of this study.

Model I is aimed at estimating the linear structural relations between the intervention variables and the outcome variables with social support as the intervening factor (see Figure 3). Model II is designed to mainly estimate the linear structural relations between the intervention variables and the outcome variables with school crime and victimization as the intervening factors (see Figure 4).

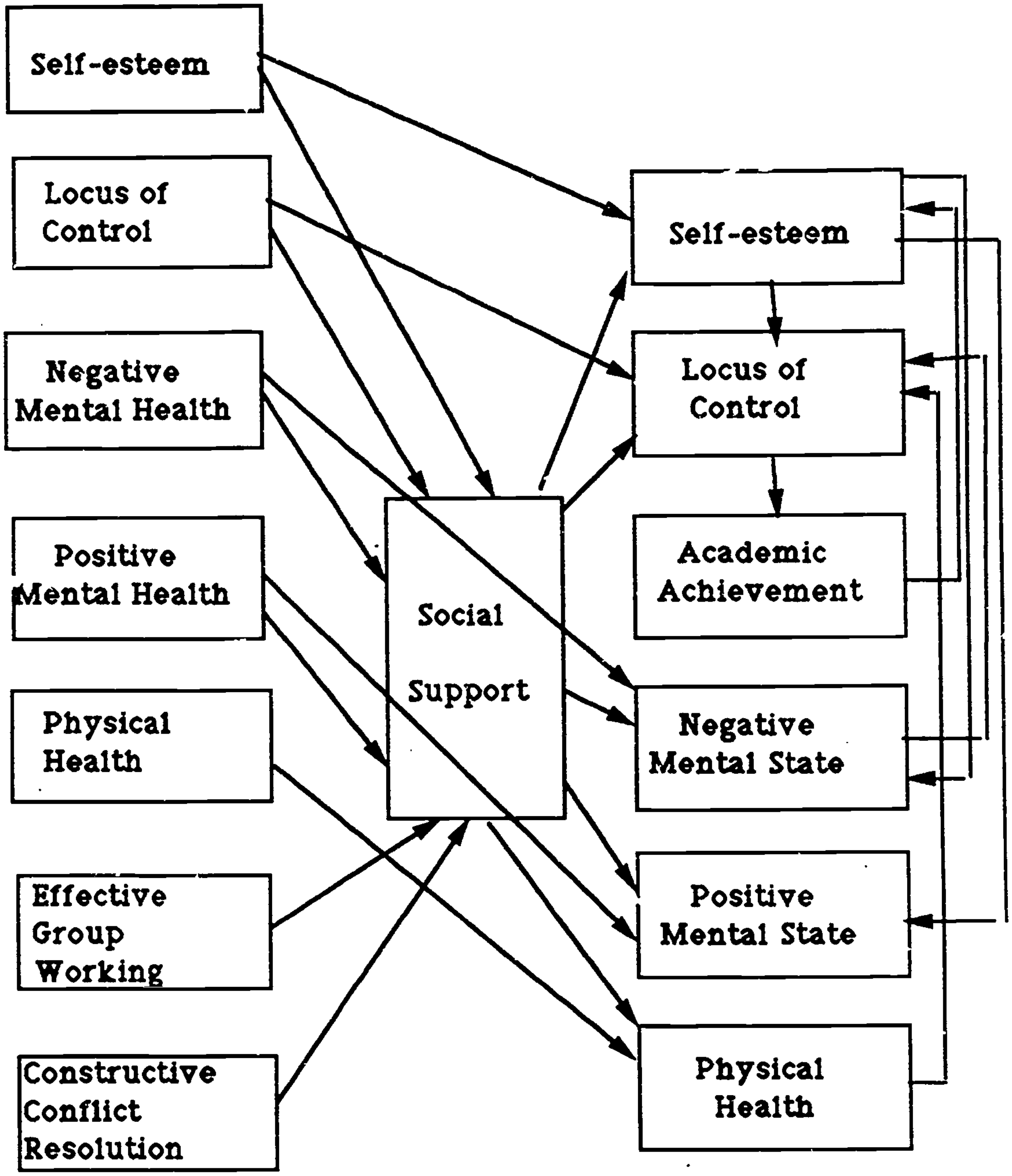
The relationships between the pre-measured variables and the intervention or climate variables will also be examined in those two models for assessing the predicted reciprocal relations as Deutsch's crude law asserted.

The computer program of LISREL VII (VAX/VMS version) will be used to analyze both Model I and Model II. Given the improved features in LISREL version 7, we hope that we can estimate the measurement model and causal model simultaneously so that we may preserve the statistical integrity of our general models.

Model III will be an estimation of the baseline of the factor structures across the three sites within the school among the intervention variables and the climate variables, which will include three separate factor structures.

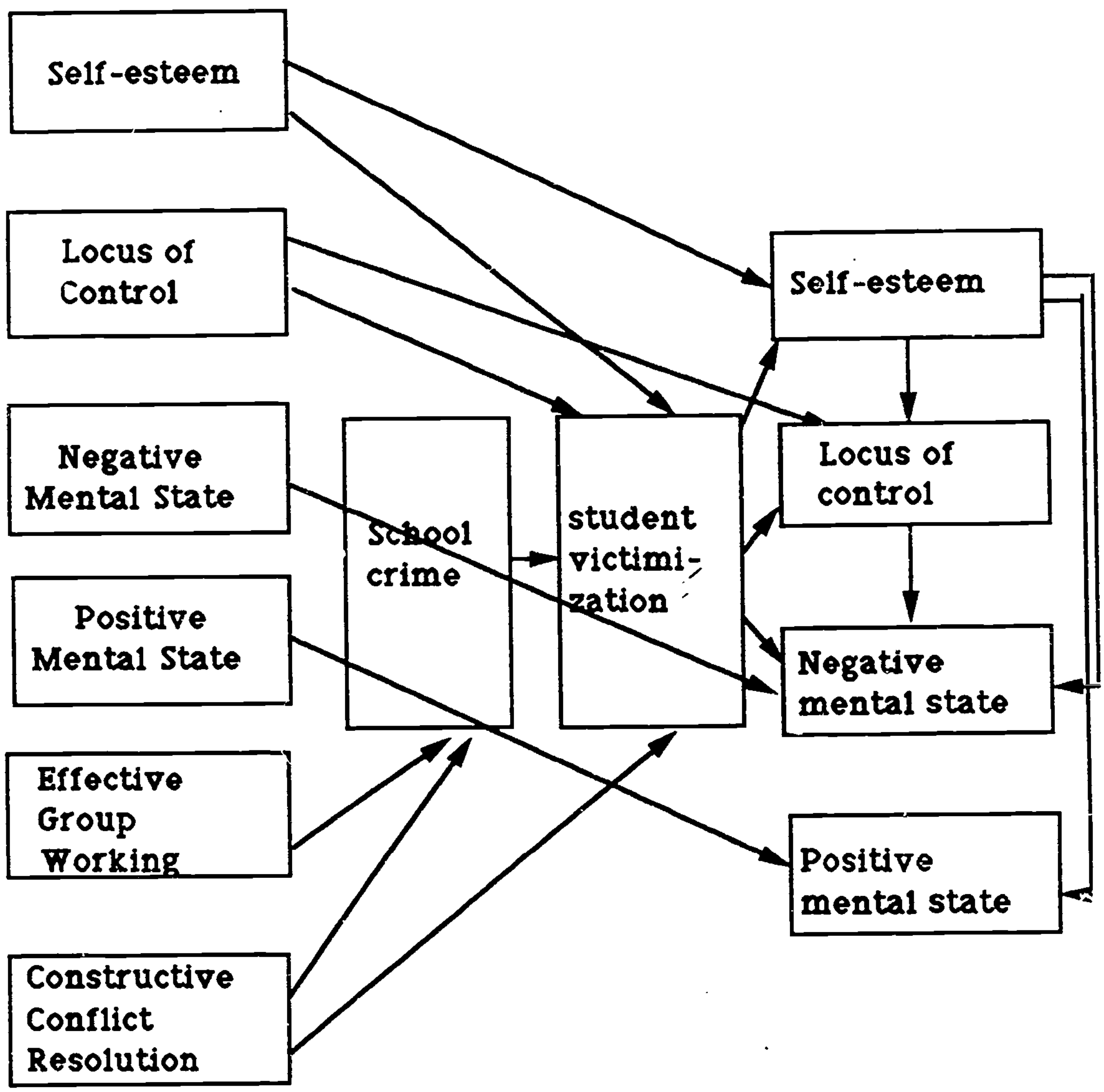
Model IV will be built to test the effects of the intervention relative to the type of program each site receives. According to the design of the intervention, the three sites within the school would receive different training programs. Site 1 would only receive conflict resolution training, Site 2 would receive training in both conflict resolution and cooperative learning, and Site 3 cooperative learning only. Based on this design, we expect that the mean structures of the variables in Model III will differ accordingly across sites, if the interventions were successfully carried out. Only post-data will be used in Model III and Model IV.

Figure 3 LISREL Design--Model I



Note Figure 3 displays the hypothesized Model I for LISREL estimation. The variables indicated are all latent. The ones on the left include five pre-measured variables; the two intervention variables on the left and all the rest are post-measured.

Figure 4 LISREL Design--Model II



Note Figure 4 displays the hypothesized Model II for LISREL estimation. The variables indicated are all latent. The ones on the left include four pre-measured variables; the two intervention variables on the left and all the rest are post-measured.

A number of issues may be encountered during the analysis, regarding the fulfillment of the statistical assumptions of LISREL. The estimation procedure which will be used is maximum likelihood (ML) method. The fitting function of this method derives from the multinormal distribution of the observed variables. When the multivariate normality assumption is fulfilled, the ML estimators are consistent and asymptotically effcient.

Consistency means that as a sample becomes arbitrarily large, the estimator of a population parameter converges on that parameter. Efficiency implies that the estimator of a parameter is a unbiased one with minimum variance. If the multinormality assumption is fulfilled, the ML estimators are consistent and asymptotically efficient. Thus the asymptotic variance-covariance matrix estimated from the model will be equal to the inverse of Fisher information matrix. The X^2 statistic will also be correct.

The violation of multinormality does not affect the consistency of the ML estimators of the population parameters. However, excessive kurtosis ("fatter" or "thinner" tails than the normal distribution) usually eliminates asymptotic efficiency; thus the estimated asymptotic variance-covariance matrix and the X^2 estimate may be inaccurate (Bollen, 1989).

Table 1 shows the summary statistics of the observed variables. The values of skewness and kurtosis for most variables are around zero, except for "Pdrugs", "Pweapo", and the last three. The last four columns indicate that these variables are heavily concentrated at the left end of the scales, which suggests that these variables may be censored below. Assuming their corresponding latent variables are normally distributed with unknown mean μ and standard deviation σ, the normal inverse transformations were conducted for those and together with other censored variables; and both polychoric correlations and canonical correlations were computed for them by PRELIS. However, the resulting overall correlation matrices failed to be positive definite. Therefore, the original untransformed covariance matrix was used for the final analysis, which may require that caution be taken in interpreting the analytical results.

Another factor which may drastically affect a chi-square estimate is sample size. Thus when the number of missing data is large, even if the normality assumption is met, the X^2 can still be quite off. Table 2 (Appendix B) contains the pattern of missing values on the observed variables we used in the models. Only 13 cases have complete information on every variable. There are 536 cases missing 30 values across the variables used. This large concentration of missing cases is due to the fact that about two thirds of the students who took the pre-test did not complete the post-test. Thus students who took only the

Table 1

Univariate Summary Statistics

VAR.	MEAN	SD	SKEWNESS	KURTOSIS	MIN.	FRQ.	MAX.	FRQ.
PFAMUP	3.425	1.255	-0.349	-0.702	1.0	37	5.0	94
PSCHUP	3.575	0.908	-0.477	0.471	1.0	11	5.0	51
PWRKUP	3.306	1.020	-0.367	-0.045	1.0	19	5.0	33
PFRIUP	3.559	0.923	-0.488	0.453	1.0	12	5.0	50
PGDQUA	3.480	0.647	-1.114	1.190	1.0	4	4.0	198
PUSELE	2.881	0.905	-0.271	-0.869	1.0	22	4.0	104
PNOGOO	3.106	0.945	-0.676	-0.648	1.0	23	4.0	157
PLUCK	3.884	1.071	-1.062	0.699	1.0	18	5.0	111
PWASTP	3.672	1.071	-0.557	-0.460	1.0	10	5.0	85
PPLANS	3.639	0.997	-0.582	-0.102	1.0	10	5.0	67
ACHIEV	44.789	9.200	-0.358	-0.299	21.0	1	66.5	1
PUPSET	2.918	1.510	0.593	-0.705	1.0	58	6.0	28
PTENSE	3.031	1.581	0.351	-1.064	1.0	69	6.0	27
PSAD	2.491	1.642	0.849	-0.530	1.0	140	6.0	28
PFELTD	2.835	1.468	0.593	-0.623	1.0	66	6.0	21
PCHEER	3.872	1.504	-0.338	-0.924	1.0	28	6.0	49
PWAKIN	3.277	1.485	0.123	-0.962	1.0	48	6.0	25
PLIFEF	3.559	1.484	0.039	-1.053	1.0	27	6.0	38
PSTABL	3.766	1.552	-0.173	-1.088	1.0	30	6.0	53
PYOUPE	4.341	1.332	-0.824	0.237	1.0	18	6.0	70
PYOUIL	4.511	1.310	-0.963	0.420	1.0	15	6.0	81
PWORR	4.313	1.605	-0.552	-0.874	1.0	22	6.0	120
GDQUAL	3.472	0.610	-0.857	0.509	1.0	5	4.0	456
USELES	2.781	0.909	-0.101	-0.959	1.0	57	4.0	216
NOGOOD	3.015	0.951	-0.509	-0.851	1.0	59	4.0	337
LUCK	3.939	0.989	-1.089	1.028	1.0	29	5.0	261
WASTEOF	3.691	1.033	-0.683	-0.103	1.0	27	5.0	185
PLANSUC	3.689	0.993	-0.696	0.193	1.0	29	5.0	173
UPSET	3.110	1.558	0.410	-1.040	1.0	115	6.0	75
TENSE	3.078	1.563	0.286	-1.063	1.0	157	6.0	63
SAD	2.774	1.739	0.573	-1.062	1.0	275	6.0	86
FELTDEP	2.901	1.501	0.462	-0.945	1.0	156	6.0	41
CHEER	3.941	1.473	-0.369	-0.[illegible]08	1.0	54	6.0	118
WAKING	3.262	1.541	0.100	-1.119	1.0	129	6.0	61
LIFEFUL	3.638	1.521	-0.082	-1.074	1.0	74	6.0	102
STABLE	3.967	1.541	-0.417	-0.959	1.0	65	6.0	137

VAR.	MEAN	SD	SKEWNESS	KURTOSIS	MIN.	FRQ.	MAX.	FRQ.
YOUPEP	4.420	1.282	-0.850	0.442	1.0	36	6.0	177
YOUILL	4.389	1.303	-0.774	0.036	1.0	32	6.0	167
WORRIED	4.237	1.607	-0.522	-0.840	1.0	65	6.0	261
PEFFGR	3.902	0.915	-0.758	0.608	1.0	7	5.0	95
POTHDO	3.254	1.189	-0.180	-0.815	1.0	29	5.0	60
POTHSH	3.677	0.946	-0.406	-0.354	1.0	4	5.0	68
PIMPGP	3.469	1.080	-0.375	-0.349	1.0	18	5.0	63
PIMPCR	3.252	1.141	-0.217	-0.508	1.0	30	5.0	54
PIMPEM	3.330	1.191	-0.328	-0.568	1.0	27	5.0	52
PIMPCT	3.327	1.194	-0.252	-0.718	1.0	29	5.0	67
PIMPJB	3.330	1.196	-0.243	-0.701	1.0	23	5.0	54
PDRUGS	1.45	0.743	1.713	2.536	1.0	237	4.0	11
PDRINK	1.541	0.746	1.265	0.969	1.0	209	4.0	7
PWEAPO	1.496	0.691	1.468	2.229	1.0	210	4.0	8
PVIOLE	1.661	0.736	0.845	0.072	1.0	171	4.0	5
PJEWEL	1.407	1.024	2.954	8.379	1.0	286	6.0	6
PHURT	1.382	0.958	2.986	8.801	1.0	284	6.0	4
PPHYS	1.331	0.968	3.226	9.876	1.0	307	6.0	4

Dictionary

Post-measures:	Lables:	Pre-measures:
PFAMUP	Family support	
PSCHUP	School support	
PWRKUP	Work/job Support	
PFRIUP	Friends support	
PGDQUA	Having good qualities	GDQUAL
PUSELE	Feel useless at times	USELES
PNOGOO	No good at all	NOGOOD
PLUCK	Good luck more important	LUCK
PWASTP	Planning is waste of time	WASTEOF
PPLANS	Make plans work out	PLANSUC
ACHIEV	Academic achievement	
PUPSET	Anxious, worried, upset past month	UPSET
PTENSE	Tense past month	TENSE
PSAD	Sad, discouraged, holpless past month	SAD
PFELTD	Depressed past month	FELTDEP
PCHEER	Cheerful past month	CHEER
PWAKIN	Waking up rested past month	WAKING
PLIFEF	Life full of interesting things past month	LIFEFUL
PSTABL	Emotionally stable past month	STABLE
PYOUPE	How much energy past month	YOUPEP
PYOUIL	Illness past month	YOUILL
PWORR	Worried about health past month	WORRIED
PEFFGR	Work more effectively in group last year	
POTHDO	Expect other members to do my share of work	
POTHSH	Put myself in the other person's shoes	
PIMPGP	Improvement in working in groups	
PIMPCR	Improvement in handling conflicts with friends and peers	
PIMPEM	Improvement in working with otner employees	
PIMPCT	Improvement in handling conflicts with family	
PIMPJB	Improvement in handling conflicts on the job	
PDRUGS	Drugs used on campus	
PDRINK	Liquor, beer, wine on campus	
PWEAPO	Weapons on campus	
PVIOLE	Violence, student fighting, vandalism	
PJEWEL	Forced to hand over money or things	
PHURT	Afraid of being physically hurt	
PPHYS	Being physically attacked	

pre-test would have at least 30 missing values which corresponds to the number of variables from the post-test used in the analysis.

Given a substantial number of missing cases and missing values in the data, the effective sample sizes for Model I and Model II are expected to be relatively small. To prevent drawing conclusions based on unreliable estimates due to sampling fluctuation, we will employ several partial models to cross-validate the results from the general models.

A series of t-tests have been conducted on the 17 variables from the pre-test for the two groups--the group who took both the pre- and post-tests and the one who took the pre-test only (See Table 3a in Appendix B). Out of the 17 tests, one test yielded a t-value of 2.56 with a 2-tail probability of .011. None of the rest of the tests is significant. Taking into account the number of the tests and the large degrees of freedom for each test, we are confident of the pre-existing equivalence between these two groups on the outcome variables.

Interventions

Instead of establishing a control group, the three sites were given different combinations of training programs as we stated above. The trainings were delivered directly to the teachers so that the intervention would be incorporated into the regular school curriculum. Thus students were indirectly exposed to the intervention through their trained teachers.

The advantages of this indirect exposure approach are obvious: 1. disturbance to the school functions may be kept at the minimum level; 2. teachers are relied upon so that they may be better motivated to apply the knowledge and skills they obtain from the training; 3. the effects of the training will not dissipate along with the graduation of the students; 4. financial burden may be reduced.

However, the main trade-off of this approach is that the measures of students' exposure to the intervention will depend upon whom a student takes classes with and how much the particular teacher of that course has learned from the training and how willingly the teacher will apply what he or she has learned to the class. Thus any measure of a student's exposure will be at best a rough index of the dosage of the intervention which the student receives. Therefore, we anticipate certain difficulties in demonstrating the association between the potential systematic changes on the outcome variables and the exposure to the intervention in our data analyses.

The trainings (see Appendix B for a discription of the planned trainings) were planned to continue for three successive years. But due to various reasons, we were unable to systematically carry out the third year trainings.

Subjects

As we have stated previously, the intervention was introduced to an inner-city alternative high school which has four sites at different boroughs of New York City. Due to administrative issues, only three sites participated in this project. Thus subjects generally consisted of all the students in the three sites during the period of the intervention.

Each site of the school had approximately 180 students and 13 to 15 teachers including a site coordinator. The average length of time a student stayed in the school before his or her graduation was one year and a half. The students at all three sites reported themselves to be predominantly Hispanic and black, and were for the most part economically disadvantaged. Many students had dropped out of school for periods ranging from six months to seven years prior to entering this school. The behavioral and academic problems these students experienced were confirmed by the teachers. In a random sampling of students (N=212), teachers indicated that 34% had been referred for counseling and educational services.

Attendance and academic performance were issues of concern. About 65% of the students attended classes on a regular basis. The irregularity of attendance at school increased difficulties in collecting data and obtaining a reliable measure of student exposure to the intervention.

Procedure

Teachers and staff personnel voluntarily participated in the site-specific trainings. Students were subsequently exposed to the intervention through their teachers.

Base-line data (N=641) were collected prior to the commencement of formal teacher trainings at each site. New students entering the school during the course of the intervention were eligible to volunteer for involvement in our programs. We collected our pre-exposure data at the beginning of each school cycle when new students came in. There were five times of such pre-exposure data collections. Together with the base-line data, there are 1053 cases of pre-exposure data. Post-exposure data collection began one year after the first base-line data were collected. There were three times of post-exposure data collection. The last data collection was completed two years after. The total cases of post-exposure data are 363; among them 17 cases have no pre-exposure data. Thus the pre-post-matched data set in fact has 346 cases; and the other 690 cases have only pre-test data. Except for a small group of students (N=96), most subjects completed our measures only twice: pre- and post-exposure to the intervention at each of their sites.

The small proportion of pre-post matched cases relative to the number of students who had taken the pre-test may give rise to the question what is the reason for this amount of loss. We have investigated this issue based on the students' retention rate across three consecutive school cycle (ten weeks per cycle). The three cycles, cycle 4, cycle 5, and cycle 6, during which a major portion of our data collection was conducted, were examined. Thus three cohort groups were formed. The average three-consecutive-cycle retention rate of the three cohort groups is .467. A number of t-tests were conducted on the grades from each of the courses between the group which attended school for all the three consecutive cycles and the group which missed the school for at least one of the three cycles. The overall results clearly indicate that the two groups are generally equal on their grades of the first cycles for all the three cohort groups, taking into account the multiple comparisons. For details, see Table 3b in Appendix B.

All the pre- or post-data collected at different times from all three sites will be combined and used in the analysis. Due to a large number of missing values, we will not be able to estimate the models based on cohorts. Thus more noise, beyond and above the measurement errors in the data, may be introduced by this procedure. Its theoretical consequences may be even more complicated to clarify. However, since this study is mainly focused on causal modeling and no cross-sectional comparisons are planned, this procedure will not affect our analysis in a fundamental manner.

Measures

Intervention measures

The intervention variables, including effective group working and constructive conflict resolution may be operationalized in two ways: constructing items which measure the amount of change on the two variables or constructing instruments which measure student exposure to the intervention. Although the former does not provide information about whether possible changes are yielded by the intervention per se, or some other sources, it is crucial to construct the measure for detecting the hypothesized causal structure. On the other hand, it is important for the evaluation of the intervention to construct exposure measure. If the two types of measures are congruent, that is, they are significantly correlated to each other, then the causal relations based on the first type of measure may be considered, to certain extent, as the actual consequences of the intervention. However, if the association is not significant, then the causal structural modeling may be only taken as theory testing; and the evaluational study will have to be conducted separately. On the other hand, LISREL Model IV is designed to provide a more efficient evaluative tool than a student exposure measure, given the anticipated noise in the latter. If differential effects as we have designed are detected by this model, we would have certain degree of confidence to believe that the interventions are effective. The degree of confidence depends on to what extent the actual effect pattern matchs the designed pattern.

An independent source of information regarding evaluation may be provided by our qualitative analysis conducted by Mitchell (1991). The results of Model IV will be cross-validated by the qualitative findings.

Built in our General Student Questionnare for post data collections, there were a number of items with a 5-point scale asking students to report whether there were any improvements in effectively learning with others in a small group and constructively solving conflicts with fellow students, family members, friends or co-workers at their work places (see Appendix A).

The construct validity for the above change variables is examined by regressing each of the variables on the subjects self-reported favorable or unfavorable reactions to the corresponding training contents (Khattri, 1991). Table 4 shows that 31 percent of variance on the variable of effective group working is explained by whether the students liked the training materials in effective group working and considered them being useful. This result indicates significant convergent validity of the variable of effective group working. On the other hand, only 13 percent of variance on the variable of constructive conflict resolution is attributable to the same independent variables, which suggests a clear distinction between the two constructs and a high discriminant validity for the variable of effective group working.

The results on Table 4 were obtained from Site 2 and

Site 3, since cooperative learning intervention was only introduced to these two sites. For checking construct validity of constructive conflict resolution, we select information only from Site 1 for the similar reason as the above. The results are shown in Table 5. 25 percent of the variance in the variable of constructive conflict resolution was explained by the items asking students to recall whether the training materials relating to conflict resolution were discussed and useful for them. Comparatively, 24 percent of the variance on the variable of effective group working was explained by the same items. Although we did not obtain as high convergent and discriminant validity as effective group working, we have enough information to believe that the variable of constructive conflict resolution does measure the improvements pertaining to the training content of conflict resolution.

One point should be mentioned, that is, the dependent variables here used are simple sum of respective items in the questionnaire, which are different to a certain degree from the latent factors estimated in our models.

Table 4

Regression Analysis of the Construct Validity of Effective Group Working[a]

Criterion Variables[b]/ Predictors[c]	ß	Sig. t	Partial R	R	R^2	F	P
Effective group working/							
Like	.45	.001	.42				
Useful	.17	N.S.	.18	.56	.31	13.39	<.001
Constructive conflict resolution/							
Like	.37	.01	.33				
Useful	-.02	N.S.	-.02	.37	.13	4.6	<.05

Note

[a]This analysis was conducted by Khattri (1991). The data were from Sites 2 and 3 where cooperative learning was implemented.

[b]The two dependent variables in this table are obtained from taking the average of the relevant items in the questionnare; thus they are not exactly the same as the factors which are used in the LISREL modeling.

[c]The two predictors are obtained from subjects' self-reported reaction to the content and methods of the training in cooperative learning.

Table 5

Regression Analysis of the Construct Validity of Constructive Conflict Resolution[a]

Criterion Variables/ Predictors[b]	ß	Sig. t	Partial R	R	R^2	F	P
Constructive conflict resolution/							
Discussed	.50	.005	.46				
Useful	-.02	N.S.	-.03	.50	.25	5.79	<.01
Effective group working/							
Discussed	.36	.05	.35				
Useful	.22	N.S.	.22	.49	.24	5.44	<.01

Note

[a]This analysis was conducted by Khattri (1991). The data were from Sites 1 where conflict resolution was implemented.

[b]The two predictors are obtained from subjects' self-reported reaction to the content and methods of the training in conflict resolution.

Student exposure measures were constructed for both types of interventions. The exposure to the intervention of effective group working was composed by utilizing the variable "percentage of time used in systematic small-group teaching in your class" in the Teacher Questionnaire and the sum of the time each student spent with different teachers in the classes he or she attended; the exposure to the intervention of constructive conflict resolution was derived from the combination of trainers' rating of the teachers in terms of their knowledge and skills learned from the trainings and similarly, the time each student spent with different teachers in the classes they attended.

Social support

Methodological studies concerning the measurement of social support (Vinokur et al., 1987; Depner et al., 1984; Thoits, 1982) have indicated that the term, social support, is a general rubric that encompasses a host of more specific definitions (Depner et al., 1984). Thus traditional methods of measurement construction may not be suitable in this study, since it would be unfeasible to attempt to cover all the aspects of social support within one measure as a section of our whole questionnaire. Depner (1984) recommended using qualitative information rather than quantitative definitions, which provided a practical alternative for the measurement of social support.

Twelve items with a 5-point scale (see Appendix A) were selected to tap the social support students received from their social interactions. These interactions involved students' relationship with their family, school, work, friends, religion, neighbors, and police etc. The dimensions of social support were not pre-determined, instead that they were defined by students themselves when they answered the questionnare.

From the results of the qualitative study, we found that students defined social support as "being taken care of", "being in favor of", "providing information or explanations", "emotional or psychological help", "being on one's side", "protection", "providing services or material facilities", and "encouragement or positive feedback". These are direct summarization of the students' words, which indicates that when the word "support" is used in a social context, the students would react with those thoughts. Thus the construct validity of social support is demonstrated by our qualitative analysis.

Two other measures of social climate, school crime and student victimization (see Appendix A), were originally used in the project of "Just Community" sponsored by Lawrence Kohlberg. School crime consists of five items with a 4-point scale, asking questions such as drugs used, stealing, violence, drinking, and weapons on campus. Victimization includs 6 items with a 6-point scale tapping physical threats, sexual threats, forced handing over money, things stolen, etc. There are no validation information so far available for these measures from independent sources.

Self-evaluation

The Rosenberg Self-esteem Scale (SES) (1965) was employed to measure students' self-worth or self-acceptance aspects of self-esteem (see appendix A). The ten items were answered on a 4-point scale, from "strongly agree" to "strongly disagree", although the original actual scoring only distinguish agree or disagree. This short scale was designed to optimize ease of administration, economy of time, unidimensionality, and face validity.

The validation information was reported by Silber and Tippett (1965), Crandall (1976), and Blascovich and Tomaka (1991). Reynolds (1988) found a correlation of .38 between SES scores and overall academic self-concept. The Rosenberg measure correlated .72 with the Lerner Self-Esteem Scale. Fleming and Courtney (1984) found that SES scores correlated .78 with general self-regard, .51 with social confidence, .35 with school abilities; and no significant correlations with gender (.10), age (.13), and work experience (.07).

Locus of control was operationalized by a six-item short scale which was adapted from High School and Beyond (Lefcourt, 1981, 1984). Levenson (1981) indicated that there were three components included in the measures of locus of control: self, powerful others, and chance. Our measure, with a 5-point scale for each item from "strongly agree" to "strongly disagree", was intended to cover all three aspects despite the small number of items (see Appendix A).

The scale was developed originally for an ad hoc function, and it was not previously documented. Thus the validation information will be provided in the present study.

Mental health and physical health

The measures for mental and physical health in this study were developed by Harold J. Dupuy for the National Center for Health Statistics, originaly named as the General Well-being Schedule (GWB) (Ware et al., 1979) (see Appendix A). It is normed on a population aged 14 and above; contains six subscales: anxiety, depression, positive well-being, self-control, vitality, and general health. The first four subscales were used by Veit and Ware (1983) to compose the Mental Health Inventory (MHI). By re-organizing the items, we obtained measures for positive mental state, negative mental state, and physical health, as suggested by Veit and Ware (1983).

The General Well-being Schedule is a structured instrument for assessing self-representations of subjective well-being. The schedule originally contains 33 items, 14 with six response options, 4 with 0-10 rating bars, and 15 self-evaluation behavioral items. Scale norms are available from a sample of 79 male and 119 female students at the University of Wisconsin, Milwaukee, in a freshman psychology class (Fazio, 1977). Internal consistency coefficient was .91 for men and .95 for women.

Convergent validity was shown such that the total score correlated .47 with an independent assessment of depression, and with the six subscales correlations ranging from .27 to .44. It correlated .69 with six other measures such as PFI, PSS, CHQ anxiety scales (.41, .40, .10) and HQ, Zung, and MMPI depression scales (.35, .28, .21). The correlation with the PFI depression scale, however, was slightly higher (.50) (Andrews and Robinson, 1991).

Achievement

Student academic achievement was measured by the Regents' Competency Test (RCT) of the New York State. RCT is designed to test the minimum proficiency of the high school students in this state in order for them to obtain their diploma. Six subject areas are included: math, reading, writing, science, global history, and American history. The test is normally administered once a year in winter. Make-up tests may be offered in spring or summer.

The raw scores of the students on RCT were used in the present study. Since students would not take RCT until they were ready to graduate, their RCT scores should be a good indicator of their academic achievement after they had been exposed to the intervention. Average scores for each student of all the test subjects were computed as our achievement measure.

Chapter Five

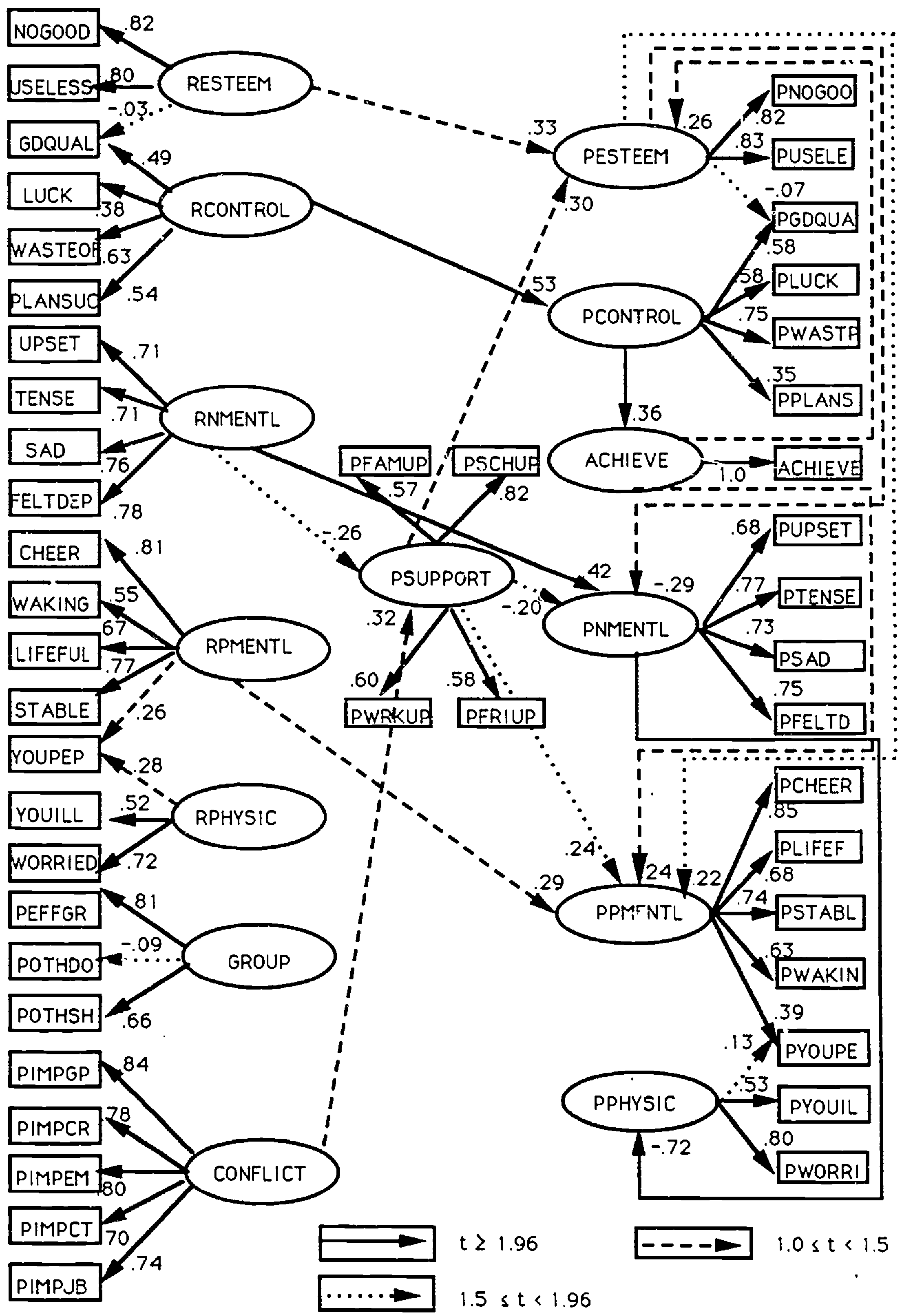
RESULTS

Model I

The main results of this study are presented in a series of LISREL diagrams. Figure 5 depicts the relationships among our target variables in Model I. The boxed items on the diagram are the manifest variables; and the circled ones are the latent variables which are measured by the manifest variables. The latent variables on the left side of the diagram are exogenous, and those on the right and the "Psupport" in the middle are endogenous.

The five exogenous variables "Resteem," "Rcontrol," "Rnmentl," "Rpmentl," and "Rphysic" are the latent factors of pre-measured variables; they stand for self-esteem, locus of control, negative mental state, positive mental state and physical health, respectively. The remaining circled variables are the latent factors of post-measured variables. They represent the two intervention variables, the intervening variable, and the six outcome variables: effective group working (Group), constructive conflict resolution (Conflict), social support (Psupport), self-esteem (Pesteem), locus of control (Pcontrol), academic achievement (Achieve), negative mental state (Pnmentl), positive mental state (Ppmentl) and physical health (Pphysic).

Figure 5 Diagram of Model I



Symbols of the Latent Variables in Model I

Latent Variables	Symbols
RESTEEM	KSI 1
RCONTROL	KSI 2
RNMENTL	KSI 3
RPMENTL	KSI 4
RPHYSIC	KSI 5
GROUP	KSI 6
CONFLICT	KSI 7
PSUPPORT	ETA 1
PESTEEM	ETA 2
PCONTROL	ETA 3
ACHIEVE	ETA 4
PNMENTL	ETA 5
PPMENTL	ETA 6
PPHYSIC	ETA 7

Note The parameter estimates between KSI's and ETA's are named as Gamma coefficients; for example, Gamma (3,2) in the tables below stands for the effect of KSI 2 on ETA 3. The parameter estimates among ETA's are named as Beta coefficients; for example, Beta (4,3) presents the effect of ETA 3 on ETA 4.

The LISREL estimation includes two parts simultaneously: the measurement structure and the causal structure. In general, the full LISREL model is defined by the three equations:

$$ETA = \beta ETA + \Gamma KSI + PSI$$

$$Y = Lambda_{y}ETA + \epsilon$$

$$X = Lambda_{x}KSI + \delta.$$

When the three structures represented by the above three equations are estimated simultaneously, LISREL produces eight output matrices of structural parameters: Lambda Y, Lambda X, Beta, Gamma, Phi, Psi, Theta Epsilon and Theta Delta. The output matrices for Model I are presented in Tables 6-10. These parameter estimates are obtained through maximum likelihood estimation.

The structural parameters between the manifest variables and the latent variables on Figure 5 are factor loadings, namely, regression coefficients. The coefficients between the manifest variables and the exogenous variables are given by Lambda-X matrix; and the ones between the manifest and the endogenous variables are given by Lambda-Y matrix. These coefficients indicate the strength of the relationships between a latent variable and its corresponding manifest variables. Numerically, an element of Lambda X or Lambda Y represents the number of units of its corresponding X_i or Y_i that is expected to change for a one-unit change in the latent variable. If more than one latent variable affects X_i or Y_i, the coefficient of Lambda

X or Lambda Y is the expected change corresponding to that latent variable while holding constant the other latent variables. A significant t level ($t \geq 1.96$) suggests that a given manifest variable is a good indicator of the latent variable. Table 6 presents the estimated Lambda coefficients with their standard errors and t-values.

There are a number of Lambda coefficients in Table 6 having a zero standard error and a zero t-value. The corresponding manifest variables are used as reference variables in order to fix an origin and a scale for each of the latent variables. Unlike in a typical regression equation where both dependent and independent variables are observed, the latent variables in LISREL do not exist in reality. Both the origin and the unit of measurement in each latent variable are arbitrary. Thus an origin and a scale must be assigned to each one of them during the estimation in order to define the latent variables. By treating the latent variables as having the form of deviation, an origin has been assigned such that each variable has a zero mean. The most useful and convenient way of assigning the units of measurement for the latent variables is to fix a non-zero value (usually one) in each column and in different rows of Lambda X and/or Lambda Y. This defines the unit of each latent variable in relation to one of the observed variables. Thus, the unit of the measurement of the latent variable will be equal to the unit of measurement of the observed variable minus its error

Table 6

ML Estimates of MODEL I (Unstandardized)

Coefficients		ML Estimates	Standard Error	T-value
LAMBDA Y				
Lambda	1,1	.969	.256	3.786
Lambda	2,1	1.000	.000	.000
Lambda	3,1	.820	.209	3.926
Lambda	4,1	.717	.188	3.805
Lambda	5,2	-.060	.146	-.414
Lambda	6,2	1.000	.000	.000
Lambda	7,2	1.034	.202	5.130
Lambda	5,3	.470	.166	2.826
Lambda	8,3	.778	.222	3.502
Lambda	9,3	1.000	.000	.000
Lambda	10,3	.435	.195	2.228
Lambda	11,4	1.000	.000	.000
Lambda	12,5	1.000	.000	.000
Lambda	13,5	1.190	.242	4.927
Lambda	14,5	1.165	.247	4.708
Lambda	15,5	1.068	.222	4.803
Lambda	16,6	1.276	.249	5.121
Lambda	17,6	.937	.227	4.126
Lambda	18,6	1.000	.000	.000
Lambda	19,6	1.150	.244	4.712
Lambda	20,6	.518	.200	2.596
Lambda	20,7	.256	.288	.887
Lambda	21,7	1.000	.000	.000
Lambda	22,7	1.850	.638	2.900
LAMBDA X				
Lambda	1,1	-.026	.155	-.165
Lambda	2,1	1.000	.000	.000
Lambda	3,1	1.071	.250	4.284
Lambda	1,2	.462	.214	2.159
Lambda	4,2	.580	.264	2.201
Lambda	5,2	1.000	.000	.000
Lambda	6,2	.831	.288	2.885
Lambda	7,3	1.000	.000	.000
Lambda	8,3	1.000	.209	4.775
Lambda	9,3	1.199	.235	5.100
Lambda	10,3	1.055	.204	5.185
Lambda	11,4	1.160	.244	4.764
Lambda	12,4	.829	.233	3.554
Lambda	13,4	1.000	.000	.000
Lambda	14,4	1.158	.249	4.653
Lambda	15,4	.327	.192	1.704
Lambda	15,5	.534	.330	1.618
Lambda	16,5	1.000	.000	.000
Lambda	17,5	1.710	.597	2.865

Coefficients		ML Estimates	Standard Error	T-value
	Lambda 18,6	1.000	.000	.000
	Lambda 19,6	-.148	.252	-.588
	Lambda 20,6	.847	.326	2.598
	Lambda 21,7	.948	.138	6.849
	Lambda 22,7	.928	.149	6.244
	Lambda 23,7	1.000	.000	.000
	Lambda 24,7	.874	.159	5.485
	Lambda 25,7	.920	.158	5.829
BETA				
	Beta 2,1	.309	.164	1.890
	Beta 3,1	.113	.209	.544
	Beta 5,1	-.275	.253	-1.089
	Beta 6,1	.327	.244	1.339
	Beta 7,1	.027	.184	.146
	Beta 3,2	.125	.241	.518
	Beta 5,2	-.402	.227	-1.770
	Beta 6,2	.292	.235	1.244
	Beta 4,3	4.134	1.848	2.236
	Beta 2,4	.021	.012	1.832
	Beta 6,4	.026	.015	1.712
	Beta 3,5	.030	.226	.133
	Beta 7,5	-.485	.202	-2.409
	Beta 3,7	.101	.330	.308
GAMMA				
	Gamma 2,1	.341	.192	1.777
	Gamma 1,2	-.017	.261	-.066
	Gamma 2,2	.114	.257	.445
	Gamma 3,2	.652	.313	2.083
	Gamma 1,3	-.171	.116	-1.474
	Gamma 5,3	.392	.157	2.492
	Gamma 1,4	.024	.134	.180
	Gamma 6,4	.285	.156	1.832
	Gamma 7,5	-.020	.197	-.101
	Gamma 1,6	.096	.196	.487
	Gamma 2,6	.006	.176	.035
	Gamma 3,6	.141	.213	.662
	Gamma 1,7	.246	.140	1.755
	Gamma 5,7	.150	.161	.931
	Gamma 6,7	-.069	.158	-.437
	Gamma 7,7	.014	.120	.117
PSI				
	Psi 1,1	.419	.143	2.929
	Psi 2,2	.334	.118	2.823
	Psi 3,3	.323	.164	1.974
	Psi 4,4	71.589	14.478	4.945
	Psi 5,5	.613	.246	2.494
	Psi 6,6	.621	.246	2.524
	Psi 7,7	.231	.146	1.582

Coefficients	ML Estimates	Standard Error	T-value
THETA EPS (TE)			
TE 1,1	1.059	.229	4.619
TE 2,2	.275	.106	2.603
TE 3,3	.671	.148	4.525
TE 4,4	.570	.124	4.607
TE 5,5	.291	.067	4.342
TE 6,6	.271	.105	2.587
TE 7,7	.307	.114	2.707
TE 8,8	.762	.173	4.405
TE 9,9	.512	.167	3.070
TE 10,10	.874	.173	5.058
TE 11,11	.500	.000	.000
TE 12,12	1.237	.275	4.502
TE 13,13	1.023	.261	3.927
TE 14,14	1.279	.301	4.249
TE 15,15	.964	.234	4.126
TE 16,16	.662	.223	2.976
TE 17,17	1.345	.287	4.681
TE 18,18	1.219	.270	4.516
TE 19,19	1.111	.268	4.140
TE 20,20	1.445	.284	5.092
TE 21,21	1.233	.276	4.477
TE 22,22	.925	.525	1.761
THETA DELTA (TD)			
TD 1,1	.288	.064	4.532
TD 2,2	.292	.121	2.412
TD 3,3	.293	.135	2.163
TD 4,4	.837	.170	4.927
TD 5,5	.649	.169	3.834
TD 6,6	.697	.159	4.373
TD 7,7	1.201	.274	4.385
TD 8,8	1.217	.277	4.396
TD 9,9	1.262	.312	4.043
TD 10,10	.887	.227	3.914
TD 11,11	.763	.237	3.225
TD 12,12	1.657	.343	4.823
TD 13,13	1.268	.289	4.384
TD 14,14	.974	.266	3.658
TD 15,15	1.336	.266	5.016
TD 16,16	1.243	.277	4.489
TD 17,17	1.247	.465	2.679
TD 18,18	.284	.201	1.417
TD 19,19	1.402	.266	5.276
TD 20,20	.507	.169	3.008
TD 21,21	.341	.092	3.696
TD 22,22	.512	.120	4.280
TD 23,23	.501	.123	4.085
TD 24,24	.726	.156	4.666
TD 25,25	.655	.145	4.519

term. The observed variable is regarded as a reference variable for the latent variable. For example, "Pschup" is used as the reference variable of the endogenous latent variable "Psupport", and "Wasteof" is the reference variable of the exogenous latent variable "Rcontrol." Since the Lambda coefficient corresponding to a reference variable is fixed, its standard error is set to zero and the t-value thus cannot be estimated.

The directions of the arrows in the diagram of Figure 5 imply that the manifest variables are determined by the latent ones. All the observed variables are standardized to have unity variances so that the Lambda X's and Lambda Y's are comparable among the manifest variables.

According to the domain theory (Nunnally, 1978), the reliability of a test item is defined by the squared correlation coefficient between the item and the true score, that is, the ratio of the true variance to the observed variance, $Var(\tau_i)/Var(X_i)$ (Lord and Novick, 1968); where τ_i is a true score, and X_i is the corresponding observed item. In the context of LISREL, the reliability is defined by the squared multiple correlation between a manifest variable and its latent variable(s). (In the circumstances where factor structure is unidimensional, the R-squares are just the squared simple correlations.) It should be noted that a squared multiple correlation in LISREL provides a lower bound for the reliability of a variable, since the R-square does not include the variance of the specific component contained in a true score. We have found that the variable "Pothdo" has an extremely low reliability of .008 (Table 7, Appendix B). Most of the reliability cofficients are around .5. Since these indices are the lower bounds of reliability and only for single items, we perceive them as sufficiently high.

However, reliability is different from validity. Examining Figure 5, we may see that the Lambda coefficient of the item "Gdqual" in the pre-measure and the Lambda coefficient of the same item "Pgdqua" in the post-measure are very small negative numbers: -.031 and -.071, although their reliability indices at .226 and .306 are not the lowest.

The negative sign of the Lambda coefficients indicates that the changes on the latent and the manifest variables are in the opposite directions. The absolute values of the coefficients suggest that the item is practically useless in measuring "Self-esteem." However, this item is a good indicator of the latent variable "locus of control." The factor loadings of "Gdqual" on "Rcontrol" is .49 and "Pgdqua" on "Pcontrol" is .582. This demonstrates that a reasonably reliable variable may not be a valid measure of the trait that it purports to measure. Another poor indicator is "Pothdo" for the latent variable "Group" with the reliability of .008 and the Lambda coefficient of -.092, which is neither reliable nor valid.

LISREL gives all the unstandardized estimates of the structural parameters along with their asymptotic standard errors and the t-values (Table 6). Standard errors are estimates of the precision of each parameter estimate. Small standard errors correspond to good precision and large standard errors to poor precision. However, what is small or large depends on the units of measurement in observed and/or latent variables and the magnitude of the parameter estimate itself. Thus t-values, which are independent of the units of measurement, are used to represent the relative magnitude of the parameter estimates.

A t-value is defined as the ratio between the parameter estimate and its standard error. Figure 5 shows all the structural coefficients with a t-value equal or greater than 1.0. (The four factor loadings of "Gdqual," "Pgdqua," "Pothdo," and "Pyoupe" have t-values below 1.0. But we used dotted lines for these four estimates in order to indicate to which latent variables they were thought to be related.). The estimates with a t-value equal or greater than 1.96 are indicated by solid lines. The ones with a t-value between 1.5 and 1.96 are indicated by dashed lines, and the ones with a t-value between 1.0 and 1.5 by dotted lines.

By estimating Theta Epsilon and Theta Delta, the LISREL program extracts the true score or latent variable for each of the manifest variables (Table 6). Since "Achieve" is a one-indicator factor and its measurement error is normally not extractable, the corresponding Theta Epsilon is pre-fixed as .5. This value was based roughly on the results from other analyses. Interestingly, within the range we tried, numbers either greater or less than .5 with an increment of .1 were not acceptable to this error term. In other words, fixing the error term to any number other than .5 would make the model fail to converge.

The magnitude of Theta Epsilon or Theta Delta is also an indication of whether a variable is a reliable measure of the latent construct(s). The size of a Psi coefficient gives the idea how efficient a structural equation is (Table 6). In other words, Psi coefficients are the errors in the equation associated with latent endogenous variables rather than the manifest variables. As in a regression analysis, the disturbance term includes those variables that influence the dependent variable but are excluded from the equation.

The arrows in Figure 5 relating the exogenous variables to the endogenous variables represent the Gamma parameters. The arrows among the endogenous variables are Beta parameters. These two parameter matrices constitute the causal structure--the primary interests of our theory testing. The model to be estimated was specified according to our prior hypotheses outlined in Figure 3. Again, only coefficients with a t-value equal or greater than 1.0 are displayed.

It is important to realize that unlike most multivariate analysis, the treatment of LISREL results should be carried out in a comparative manner and with sufficient caution. Most of the LISREL estimation procedures are based on large sample normal assumptions. Although LISREL 7 has included distribution free estimation procedures such as Generalized Weighted Least Squares (GLS) and Diagonally Weighted Least Squares (DWLS), it was found that the estimation, using the distribution free methods, may not converge to the true parameter until the sample size increases to 3500 to 5000 (Bentler, 1991).

Due to missing data, our effective sample size for Model I is only 57. Given this sample size, the conventional critical value .05 for testing the significance of the parameter estimates may not be appropriate with respect to the resulting small test power. Therefore, we have kept those coefficients with a t-value equal or greater than 1.0 in our diagram; further validation is needed by a power analysis and by estimating partial models.

The major findings on Figure 5 can be tentatively summarized as follows: (a) constructive conflict resolution may have positive effects on social support among students (.32); (b) social support may have positive effects on a student's self-esteem (.30); (c) social support may reduce a student's negative mental state (-.20); (d) social support may enhance a student's positive mental state (.24); (e) a student's self-esteem may reduce his/her negative mental state (-.29) and enhance his/her positive mental state (.22); (f) a student's negative mental state may jeopardize his/her physical health indicated with a very large coefficient of -.72; (g) internal locus of control contributes to a student's academic achievement again indicated with a substantial coefficient of .36; (h) a student's academic achievement may enhance his/her self-esteem (.26); and (i) the academic achievement also promotes positive mental state (.24).

The Gamma parameters between the same latent variables but obtained at different times are viewed as stability coefficients. We found that self-evaluation and mental health are rather stable from the pre- to the post-measure. Among them, locus of control is the most stable trait across time with a coefficient .53. Surprisingly, physical health is not as stable as we would normally expect. Possible explanations for this will be explored later. The negative mental state measured on the pre-test is found to have a negative impact (-.26) on the social support measured on the post-test.

A number of power analyses were conducted to investigate the probability of rejecting the null hypotheses in regard to the parameter estimates which had a t-value between 1.0 and 1.96, given the true effects as estimated had existed (Table 8, Appendix B). The test regarding the coefficient from "Conflict" to "Psupport" (Gamma (1,7)) has a power of .39. The test regarding the coefficient from "Psupport" to "Pnmentl" (Beta (5,1)) has a power of .23. The test of "Psupport" to "Ppmentl" (Beta (6,1)) has a power of .33. The test of "Pesteem" to "Pnmentl" (Beta (5,2)) has a power of .48. The test of "Pesteem" to "Ppmentl" (Beta

(6,2)) has a power of .28. The test of "Achieve" to "Pesteem" (Beta (2,4)) has a power of .34. The test of "Achieve" to "Ppmentl" (Beta (6,4)) has a power of .52. The power of "Rnmentl" to "Psupport" (Gamma (1,3)) is .28. The power for "Resteem" to "Pesteem" (Gamma (2,2)) is .07. And the power for "Rpmentl" to "Ppmentl" (Gamma (6,4)) is .53. The power analysis for the coefficient from "Psupport" to "Pesteem" (Beta (2,1)) was not able to be completed. The non-centrality parameter which is required for the analysis was not obtainable, due to the model identification problem resulted from fixing this coefficient to zero. However, this is an indication that a serious specification problem may occur if this parameter is omitted from the model.

Obviously, this sample did not provide enough power for the tests of the above effects. It is possible to assume that given a larger sample size, these effects could have produced t-values greater than the conventional significance level. Furthermore, other coefficients omitted from the diagram could also have been significant. Bollen (1989) has referred to a sample of 138 in one of his studies as having a moderate size.

To test how large a sample is needed to be in order to achieve a sufficient test power in our case, holding other relevant factors such as effect sizes and standard deviations constant, we set the NO (number of observations) parameter equal to 200 in the LISREL command lines and re-ran the same model. The results showed that the smallest t-value for the eleven parameters is 2.053; and most of them are greater than 3.

However, the assumption that the sampling fluctuation does not significantly affect the standard deviations or the covariances may not be tenable. To draw safe conclusions from Model I, we need further empirical evidence by increasing the effective sample size from the same subject pool. Such a study requires a reduction in the number of variables in the analysis. This validation study was done by analyzing several partial models which are presented later.

Figure 5 depicts only direct effects between the latent variables. However, since the variables are inter-related, their effects may reach each other through many indirect paths. Thus, an examination of the total effects of one variable on another will give a more completed picture of the relationships among them. Table 9 presents the total effects of ETA on ETA and KSI on ETA, that is, the effects among the endogenous variables and the effects between the exogenous and endogenous variables.

The total effects are more detailed results compared to the diagram of Figure 5. It should be mentioned that Table 9 presents unstandardized coefficients. If we compare the row with the column of "Achieve," we will find that all the estimated coefficients on the row are much greater than those on the column except for the effect of "Achieve" on itself. This is so because "Achieve" has a very large mean

(44.79) and variance (84.64) (see Table 1), the effects from "Achieve" will be much smaller than the effects toward "Achieve". That is, one unit of change in "Achieve" would only cause a slight difference in its measure in an affected variable, while a one-unit change of any variable would reflect as a large change in its measure of "Achieve." However, this discrepancy only reflects the scaling differences between "Achieve" and other variables.

Table 9 shows that "Psupport" has effects on "Pesteem" (.324), "Pnmentl" (-.405), "Ppmentl" (.439), and "Pphysic" (.224); "Pesteem" has effects on "Pnmentl" (-.407), "Ppmentl" (.31), and "Pphysic" (.197); "Pcontrol" has effects on "Achieve" (4.182), and "Ppmentl" (.137); "Achieve" has effects on "Pesteem" (.021), "Pnmentl" (-.009), "Ppmentl" (.033), and "Pphysic" (.004); "Ppmentl" has a rather large effect on "Pphysic" (-.486); and the intervention variable "Conflict" affects both "Psupport" (.246) and "Pesteem" (.08). The ratio of these estimates to their corresponding standard errors are all greater than one.

Table 10 shows the maximum likelihood estimates of the correlation matrix of the exogenous variables. By definition, the causes for the exogenous variables are outside our research interests. Thus, the relationships among these variables are not viewed in the causal context. However, a significant association between two exogenous variables does reveal that (a) when one of these two variables is demonstrated to be a cause of a variable, the effect of that variable contains some contributions from the other; (b) a causal relationship may exist between these exogenous variables if a theoretical justification is available, although the causal relationship is not tested.

Phi (7,6) indicates that the two intervention variables "Group" and "Conflict" are correlated with a correlation estimate of .392. The largest correlation estimate is -.765 for Phi(5,3), which is the correlation between "Rphysic" and "Rnmentl."

It is also useful to examine the unstandardized PHI matrix, since it gives variance estimates for each of the exogenous variables. Notably, all the variance estimates have a t-value greater than 2.0 except for Phi (5,5) which is the variance of "Rphysic" with a variance estimate of .456. Correspondingly, the variance of "Pphysic" is .49 (Table 1) which is also the smallest among the variances of endogenous variables. There is no standard errors and t-values available for the variances of the endogenous variables. The small variance on both the pre- and post-variable of physical health, to certain extent, may explain why its stability coefficient is not as significant as we might have expected.

Table 9 Total Effects for Model I

Total effects of ETA on ETA

	Psupport	Pesteem	Pcontrol	Achieve	Pnmentl	Ppmentl	Pphys
Psupport	.000	.000	.000	.000	.000	.000	.000
Pesteem	.324	.012	.088	.021	-.002	.000	.009
Pcontrol	.164	.134	.012	.003	-.019	.000	.103
Achieve	.680	.556	4.182	.012	-.080	.000	.424
Pnmentl	-.405	-.407	-.036	-.009	.001	.000	-.004
Ppmentl	.439	.310	.137	.033	-.003	.000	.014
Pphysic	.224	.197	.017	.004	-.486	.000	.002

Standard errors for the total effects of ETA on ETA

	Psupport	Pesteem	Pcontrol	Achieve	Pnmentl	Ppmentl	Pphys
Psupport	.000	.000	.000	.000	.000	.000	.000
Pesteem	.167	.019	.060	.012	.013	.000	.030
Pcontrol	.190	.234	.019	.005	.144	.000	.334
Achieve	.833	.953	1.876	.019	.595	.000	1.389
Pnmentl	.250	.230	.031	.007	.005	.000	.012
Ppmentl	.238	.238	.086	.015	.019	.000	.046
Pphysic	.202	.130	.016	.003	.202	.000	.006

Total effects of KSI on ETA

	Resteem	Rcontrol	Rnmentl	Rpmentl	Rphysic	Group	Conf
Psupport	.000	-.017	-.171	.024	.000	.096	.246
Pesteem	.345	.168	-.056	.008	.000	.050	.080
Pcontrol	.046	.672	-.036	.004	-.002	.159	.039
Achieve	.189	2.779	-.148	.016	-.008	.657	.162
Pnmentl	-.139	-.063	.461	-.010	.000	-.046	.050
Ppmentl	.106	.117	-.076	.295	.000	.063	.039
Pphysic	.067	.030	-.228	.005	-.020	.025	-.004

Standard errors for the effects of KSI on ETA

	Resteem	Rcontrol	Rnmentl	Rpmentl	Rphysic	Group	Conf
Psupport	.000	.261	.116	.134	.000	.196	.140
Pesteem	.195	.273	.047	.044	.002	.183	.060
Pcontrol	.086	.314	.064	.023	.021	.216	.061
Achieve	.352	1.739	.273	.095	.088	.928	.260
Pnmentl	.108	.151	.164	.055	.001	.104	.154
Ppmentl	.100	.157	.065	.164	.003	.105	.153
Pphysic	.057	.079	.113	.030	.197	.057	.121

Table 10

ML Estimates of Variances and Covariances of Exogenous Variables for Model I

Coefficients (PHI)	ML Estimates (Unstandardized)	ML Estimates (Standardized)	Standard Error	T-value
Phi 2,1	.234	.495	.108	2.167
Phi 3,1	-.341	-.422	.153	-2.232
Phi 4,1	.241	.322	.136	1.774
Phi 5,1	.153	.311	.107	1.440
Phi 6,1	.107	.198	.097	1.098
Phi 7,1	.115	.165	.113	1.018
Phi 3,2	-.204	-.284	.141	-1.439
Phi 4,2	.269	.407	.141	1.911
Phi 5,2	.137	.315	.105	1.313
Phi 6,2	.177	.373	.105	1.693
Phi 7,2	.252	.406	.125	2.008
Phi 4,3	-.418	-.369	.209	-2.004
Phi 5,3	-.572	-.765	.231	-2.477
Phi 6,3	-.123	-.151	.142	-.869
Phi 7,3	-.159	-.149	.168	-.942
Phi 5,4	.189	.273	.147	1.288
Phi 6,4	.181	.241	.136	1.330
Phi 7,4	.246	.251	.163	1.509
Phi 6,5	.075	.151	.103	.732
Phi 7,5	.070	.109	.120	.587
Phi 7,6	.276	.392	.129	2.137
Phi 1,1	.534		.182	2.940
Phi 2,2	.418		.198	2.109
Phi 3,3	1.227		.427	2.871
Phi 4,4	1.046		.400	2.614
Phi 5,5	.456		.263	1.731
Phi 6,6	.542		.243	2.232
Phi 7,7	.918		.263	3.494

ML Estimates of Variances of Endogenous Variables (Unstandardized)

Variables	ML Estimates
Psupport	.549
Pesteem	.579
Pcontrol	.643
Achieve	84.262
Pnmentl	1.080
Ppmentl	1.023
Pphysic	.490

The main findings from Model I is summarized in Figure 5. As we have stated, the model specification was based on the theoretical hypotheses derived from previous research. However, given the size of our model, which has 14 latent variables (correspondingly 47 manifest variables), numerous empirical relationships may potentially exist in our data, due to either the true relationships or sample fluctuation. Thus, many alternative models may be equally possible; in particular, these alternative models may produce the same model fit.

If three Y-variables exist, there are at least 10 possible patterns (Figure 6, Appendix C): three models with Y_1 as the common cause, three with Y_2 as the common cause, three with Y_3 as the common cause, and a full non-recursive model (Joreskog and Sorbom, 1989). With four endogenous variables, the permutation is 24; counting the non-recursive ones, there will be many more. Therefore, with a large model, often it may be impossible to estimate all the potential relationships among the variables. Because, first of all, there rarely exist theoretical justifications which give unequivocal guidelines directing the model specification; secondly, even when such guidelines are available, the empirical situation may not allow all the parameters to be estimated, due to the constraints of model identification.

We have followed three policies in our analysis: (1) specify and test only those parameters about which we have derived specific hypotheses according to prior knowledge; (2) fix those parameters which do not reveal meaningful theoretical implications in our theory testing. For example, the relationships between negative mental state and positive mental state are not within our main research interests, since there is no great uncertainty about their relationships. As a consequence of pre-fixing certain parameters, the overall model fit will be reduced. (3) For those parameters which are set free to be estimated, alternative models are tested if they are equally plausible or have been suggested by previous research, in order to ascertain whether the data support our hypotheses better. For example, the relationships between self-esteem and achievement may have three possible patterns: a reciprocal relationship, a uni-directional relationship which specifies self-esteem as a cause of achievement, and a uni-directional relationship which specifies achievement as a cause of self-esteem. Given our hypothesis that achievement is a cause of self-esteem, we verify our claim by assessing the goodness of model fit resulting from each of the three alternative models.

LISREL produces four measures of overall goodness of fit:

(a) Chi-square (X^2)

(b) Goodness of fit index (GFI)

(c) Adjusted goodness of fit index (AGFI)

(d) Root mean squared residual (RMR)

In general, ML minimizes the fit function

$$F_{ML} = \log|\Sigma| + tr(S\Sigma^{-1}) - \log|S| - (q+p).$$

Where for a square matrix A, $|A|$ denotes the determinant of A and tr(A) denotes the sum of the diagonal elements of A. Σ is the covariance matrix produced under the specified model; and S is the sample covariance matrix. The q and p in the last term of the equation are the number of x and y variables. It is clear that if Σ equals S, then, $\log|\Sigma|-\log|S|$ becomes zero and $tr(S\Sigma^{-1})$ equals (q+p). Thus, F_{ML} and X^2, which is (N-1) times the minimum of the fit function F_{ML}, both are zero with a p-value of 1, which indicates a perfect fit between the two covariance matrices. The greater the F and X^2 are, the smaller the p-value is, and the poorer the fit is between the proposed model and the data.

Comparing the three X^2's associated with the three alternative models of achievement and self-esteem, holding other relationships constant, we found that a X^2 with 980 degrees of freedom associated with the hypothesized model was 500.07. The test for the alternative model having self-esteem as the cause of achievement yielded a X^2 of 502.60 with 980 degrees of freedom. The reciprocal model had a X^2 of 497.50 with 979 degrees of freedom. However, the estimation of the reciprocal model produced a negative estimate for the R-square of the structural equation for "Achieve." This indicates possible misspecifications.

Since there is no degree of freedom for testing the alternative uni-directional model, the difference of X^2's associated with the two models are not statistically testable. But the hypothesized model has a slightly smaller X^2. Although the estimation of the reciprocal model is converged, the estimated negative R-square prevents us from making a conclusion regarding this test. Thus, the hypothesized model is considered to be tenable.

However, given a lack of strong evidence favoring the hypothesized model and the fact that the three models all have theoretical bases, as we presented in Chapter Three, the possible existence of the other two models cannot be eliminated.

Modification indices were utilized for improving the model fit. However, when modification indices are used, LISREL is no longer purely confirmatory, since the modification procedure is data driven. Thus, modification information must be used with prudence, and theoretical considerations should take priority. The utilization of modification information may be seen by comparing the designed model (Figure 3) with the result model (Figure 5). The direct linkages between "Achieve" and "Ppmentl' and between "Pnmentl" and "Pphysic" do not exist in the designed model. However, the modification indices showed large values for these two parameters, which implied that if these values were released for estimation, the sizes of X^2-differences might be as large as the modification indices, if the statistical assumptions were fulfilled.

A relatively large modification index also occurred for the parameter Beta (4,1), from "Psupport" to "Achieve." However, the coefficient estimated by releasing the parameter was negative, which was not justifable on theoretical grounds and the way our survey questions were constructed (see Appendix A). Thus the parameter Beta (4,1) remained fixed.

The fit function for ML is derived from the maximum likelihood principle based on the assumption that the observed variables have a normal distribution (Joreskog, 1967; 1989). Under the assumption of multivariate normality, ML are optimal in the sense of being most precise in large samples compared to other fit functions. However, if the distributions deviate largely from normality, the chi-square measure and standard errors of parameter estimates may not be correct.

Another problem with using chi-square is that the sample size must be sufficiently large. Research has suggested that in small samples the chi-square statistic tends to be too large, which leads to too frequent of rejections of Ho (Bollen, 1989) ("Ho" is uaually related to the more parsimoneous model.). On the other hand, small sample size reduces the power of a test. As we showed before, the estimator of chi-square is $(N-1)F_{ML}$, thus for the same value of F_{ML}, the chi-square as well as the power of the test increases as N increases. This is a paradoxical situation.

Furthermore, a small sample size leads to large standard errors estimated for each free parameter, which reduces the t statistics. Therefore, the influence of sample size on chi-square is rather complicated. However, large sample sizes and the departures from normality generally tend to increase X^2 over and above what can be expected due to specification errors in the model (Joreskog and Sorbom, 1989). Given the sample size of 57, the chi-square statistic can be treated only in a comparative manner rather than as an absolute test statistic, even if all the basic assumptions are fulfilled. The chi-squares associated with the differences (X^2-difference) between alternative models are fairly reliable and informative.

The chi-square measure is also dependent on what type of matrix is analyzed. If other conditions are fulfilled, a chi-square statistic is correct for ULS (Unweighted Least Square), ML, GLS, and DWLS, if a covariance matrix is analyzed. But if a correlation matrix is analyzed with ML, X^2 is correct only if the model is scale-invariant and the condition, $diag(\Sigma)=diag(S)$, holds.

We used the covariance matrix for Model I and the correlation matrix for Model II. Thus, the above two conditions had to be examined for Model II. According to Joreskog (1989), the fit functions for ML, GLS, WLS, and DWLS have the property of scale invariance. Since ML was used for our analysis, the first condition was fulfilled. For the second condition, a convenient way is to check the diagonal of the fitted residuals (the differences between the covariance matrix produced from the model and the sample covariance matrix) in the output. We found that only two elements of fitted residuals on the diagonal were different from zero, -.028 and .003. Thus we believe that our results from analyzing the correlation matrix are no different than the results from analyzing the covariance matrix for Model II.

The other two fit indices are less subject to the influence of sample sizes, however, their distributions are unknown. The values of these two measures should be between zero and one. The closer they are to one, the better the fit is.

The root mean squared residual (RMR) is a measure of the average of the fitted residuals and can be interpreted only in relation to the sizes of the observed variances and covariances in S.

Table 11 displays the four values of fit indices. GFI and AGFI indicate a moderate fit of the model to the data. This is what we expected because of the way the model is specified. But, the value of chi-square corresponds to a p-value of one, which implies a perfect fit. This results from the influence of the small sample size; the p-value is, in fact, overly inflated. By convention, a X^2 with its p-value equal or greater than .05 suggests an acceptable fit.

The RMR is .273. Given the large variance of the achievement measure, this value of RMR may have been pushed up. Checking the fitted residuals, we can find that all the elements greater than 1, on and off the diagonal, are associated with the achievement measure. It indicates that the RMR is to some extent inflated by the measure of achievement.

As Joreskog and Sorbom (1989) pointed out, the measures X^2, GFI, AGFI, and RMR are measures of the overall fit of the model to the data and do not express the quality of the model judged by any other internal or external criteria. It can happen that the overall fit of the model is very good while one or more relationships in the model are very poorly determined, as judged by the squared multiple correlations, or vice versa. Furthermore, if any of the overall measures indicates that the model does not fit the data well, it does not tell what is wrong with the model or which part of the model is wrong. Particularly, when the measurement model is estimated simultaneously with the causal model, the overall fit measures do not differentiate the adequacy of the goodness of fit of the two parts. Therefore, other sources of information should be considered in judging the quality of the model.

Table 11

Squared Multiple Correlation Coefficients for Structural Equations of Model I[a]

Latent Endogenous Variables	R-square
Psupport	.238
Pesteem	.422
Pcontrol	.497
Achieve	.150
Pnmentl	.433
Ppmentl	.393
Pphysic	.529

Total Coefficient of Determination for the Structural Equations[b] .... .739

Model Fit Assessment

Fit Function	Fit Value	P-value
Chi-square with df 981	500.070	1.00
Goodness of fit index	.768	
Adjusted GFI	.733	
Root mean square residual	.273	

Note

[a]These values represent the variances explained by each of the structural equations as specified.

[b]This value represents the total variance explained by the entire structural model as specified.

Table 11 gives also other measures of model assessment, which reflect the efficiency of the model specification. The squared multiple correlation coefficients for structural equations represent the variances explained by each of the structural equations. The largest figure is .529 for "Pphysic" and .497 for "Pcontrol." From Figure 5, we can see that "Pnmentl" has an effect of -.72 on "Pphysic," and "Rcontrol" has an effect of .526 on "Pcontrol".

The next two variable effectively explained by their equations are "Pnmental" and "Pesteem" with coefficients of .433 and .422. The major direct explanatory variables for "Pnmentl" are "Psupport" (-.196), "Pesteem" (-.294), and "Rnmentl" (.418). "Pesteem" is the pivotal outcome variable which passes on the effects of intervention and social support to both mental health variables, and through mental health variables to physical health, as indicated in Figure 5. The main direct contributors to the change of "Pesteem" are "Psupport" (.302), "Achieve" (.255), and "Resteem" (.327). Noticably, the magnitude of the effect of "Psupport" on "Pesteem" is almost as large as of its pre-variable "Resteem."

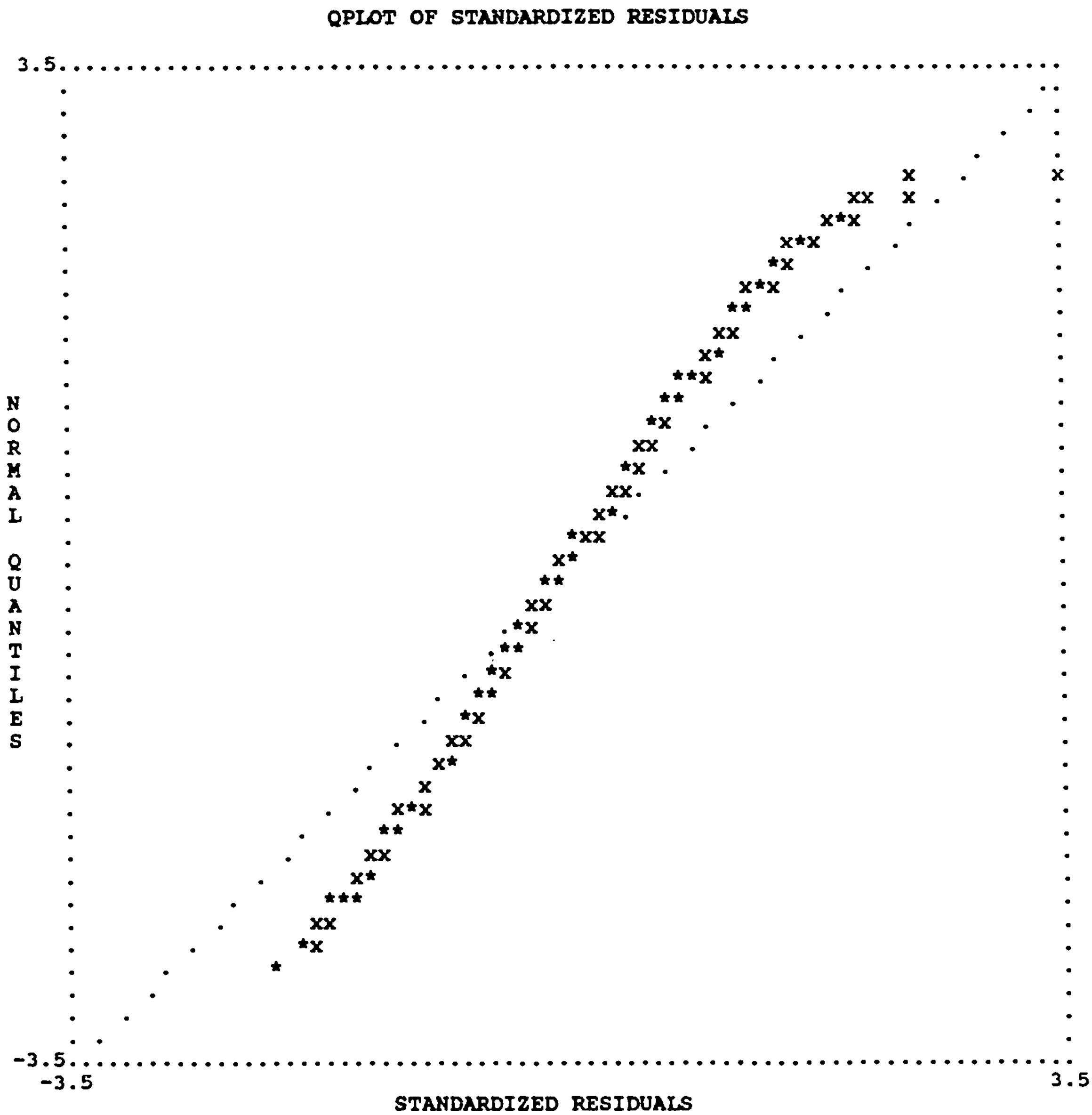
24 Percent of variance of our hypothesized intervening variable "Psupport" is explained by its equation. Figure 5 indicates that the two main direct contributors are "Conflict" and "Rnmentl" with coefficients of .319 and -.255, respectively. 15 percent of variance on academic achievement is explained mainly by internal locus of control

with a coefficient of .361.

The total coefficient of determination for the structural equations in Table 11 gives an overall percentage of variance of all the endogenous variables explained by the whole model. The coefficient is .739, which is fairly impressive. It should be noted that this measure tends to be larger when a model contains more equations (Bollen, 1989). We also noticed that this figure is larger than each one of the multiple squared correlations for the equations. It is an indication that the coefficient of determination is somewhat inflated. This is probably also due to the shared variance among the equations.

The last LISREL output measure we want to present regarding Model I, which is more relevant particularly in our case, is the Q-plot of the standardized residuals shown in Figure 7. The Q-plot is a plot of the standardized residuals against normal quantiles. Each single point is represented by an x and multiple points by an *. The 45-degree is given by small dots. If the slope of the plotted line is greater than one, that is, the angle of the line is greater than 45 degree, it is an indication of a good fit.

Figure 7 Q-Plot of MOdel I

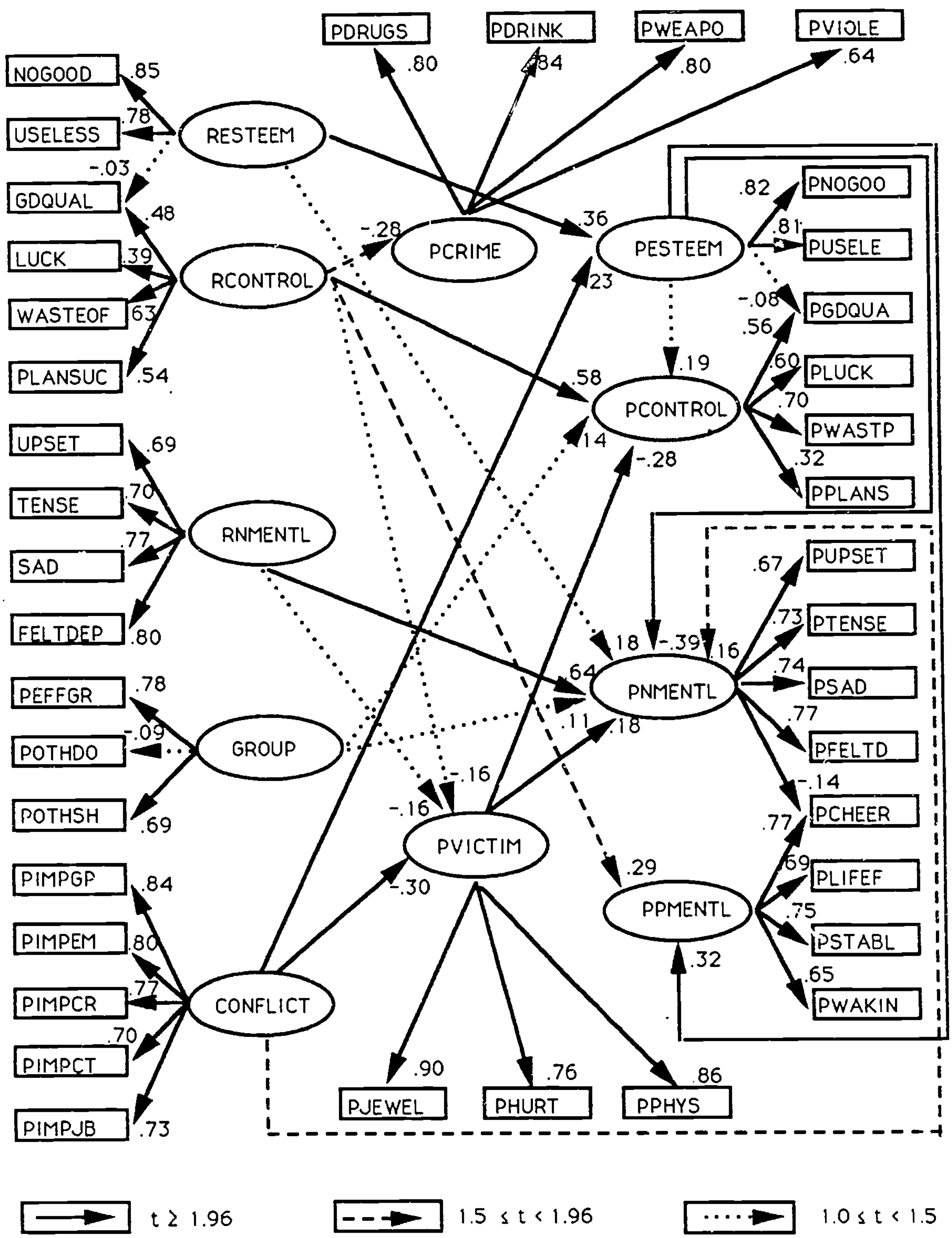


Slopes close to one correspond to moderate fits and slopes smaller than one to poor fits. Non-linearity in the plotted points are indicative of specification errors in the model or of departures from linearity or normality (Joreskog and Sorbom, 1989). Figure 7 indicated a fairly neat straight line with a slope greater than one, which suggests that there is no serious misspecification in our model and no departure from normality in our data with Model I.

Model II

Figure 8 presents the LISREL diagram of Model II. As a parallel model of Model I, school crime and a student's victimization were treated as the intervening variables rather than social support. In order to reduce the number of missing cases, we excluded the achievement variable from Model II. The academic achievement variable was measured by RCT, which came from a source independent of our data collection. Matching the RCT data in Model I to our variables drastically reduced the effective sample size for the analysis. Matching the pre- and post-data also noticably reduced the effective sample size. Thus we excluded "Rpmentl" and "Rphysic" from Model II. In order to obtain better estimation of the model, "Pphysic" was also eliminated from the analysis. As a result, the sample size for Model II reached 151. This is still not large enough to provide an adequate power for the model testing.

Figure 8 Diagram of Model II



Symbols of the Latent Variables in Model II

Latent Variables	Symbols
RESTEEM	KSI 1
RCONTROL	KSI 2
RNMENTL	KSI 3
GROUP	KSI 4
CONFLICT	KSI 5
PCRIME	ETA 1
PVICTIM	ETA 2
PESTEEM	ETA 3
PCONTROL	ETA 4
PNMENTL	ETA 5
PPMENTL	ETA 6

Figure 8 indicated that the measures for "Pcrime" and "Pvictim" are fairly adequate. Their factor loadings and reliability coefficients are of sufficient size (Tables 12 and 13, Appendix B). Compared to Table 6, the reliability estimates of the two estimations are rather close. Of course, since the the variables involved and the model specifications are somewhat different, the reliability coefficients are expected to differ accordingly.

Examining Figure 8, we may summrize our findings as follows: (a) constructive conflict resolution may increase a student's self-esteem (.23); (b) constructive conflict resolution may reduce a student's victimization (-.30); (c) constructive conflict resolution may potentially reduce a student's negative mental state (-.16); (d) effective group working may potentially have an impact on a student's locus of control (.14); (e) effective group working may potentially increase a student's negative mental state (.11); (f) a student's victimization may change his/her internal locus of control to external locus of control (-.28); (g) a student's victimization may cause his/her negative mental state (.18); (h) self-esteem may potentially lead to internal locus of control; (i) self-esteem may reduce a student's negative mental state and enhance his/her positive mental state.

Gamma matrix is free by default in LISREL. For Model I, we fixed the matrix and estimated only those parameters we had hypothesized. In order to explore possible empirical relationships, which might be suggested by the data, we used the default format for Model II to estimate all the elements of Gamma matrix. This makes Model II a combination of confirmatory and exploratory methods.

Notably, previous negative mental state may potentially reduce victimization (-.16); and previous self-esteem and effective group working (.11) may potentially increase negative mental state (.18). Although the effect sizes are small and insignificant, these results were unexpected.

The stability coefficients are all significant, especially for negative mental state (.64). It is rather heuristic that when the model included school crime and victimization, the negative mental state variable became more stable and salient. We may see from Figure 8 that there are six arrows pointed toward "Pnmentl", which may be an indication of the importance of this variable in association with school crime and victimization. Of course, the change may be due to the use of a different sample in the analysis.

Compared to Model I, Model II has a slightly better fit judging by GFI, AGFI, and RMR (.768, .733, and .273 vs. .811, .773, and .058) (Table 16). However, the chi-square value with 652 degrees of freedom is significant with three decimal digits. It is worth noting how much the chi-square statistic fluctuates with the sample size. RMR has a small value of .058. This is due to the fact that the input matrix for Model II is a correlaiton matrix, and thus, the

Table 16

Squared Multiple Correlation Coefficients for Structural Equations

Latent Endogenous Variables	R-square
Pcrime	.100
Pvictim	.187
Pesteem	.283
Pcontrol	.668
Pnmentl	.532
Ppmentl	.259

Total Coefficient of Determination for the Structural Equations.......................... .830

Model Fit Assessment

Fit Function	Fit Value	P-value
Chi-square with df 652	789.570	0.000
Goodness of fit index	.811	
Adjusted GFI	.773	
Root mean square residual	.058	

fitted residuals are smaller as compared to the case of Model I in which a covariance matrix is analyzed.

The squared multiple correlation coefficients for structural equations indicate that the equation for "Pcrime" has a relatively small R-square, .100. Checking the reason for that, we found that a mistake was made confusing the school level variable "Pcrime" with the individual level variables in the analysis. This confusion partly explains the reason why the standard deviations of the indicator variables for "Pcrime" are low, since these variables are not expected to vary systematically from one individual to another.

The determination coefficient is very high (.83) for Model II, which is another indication of the efficiency of the model. The Q-plot (Figure 9, Appendix C) suggests a moderate fit for this model.

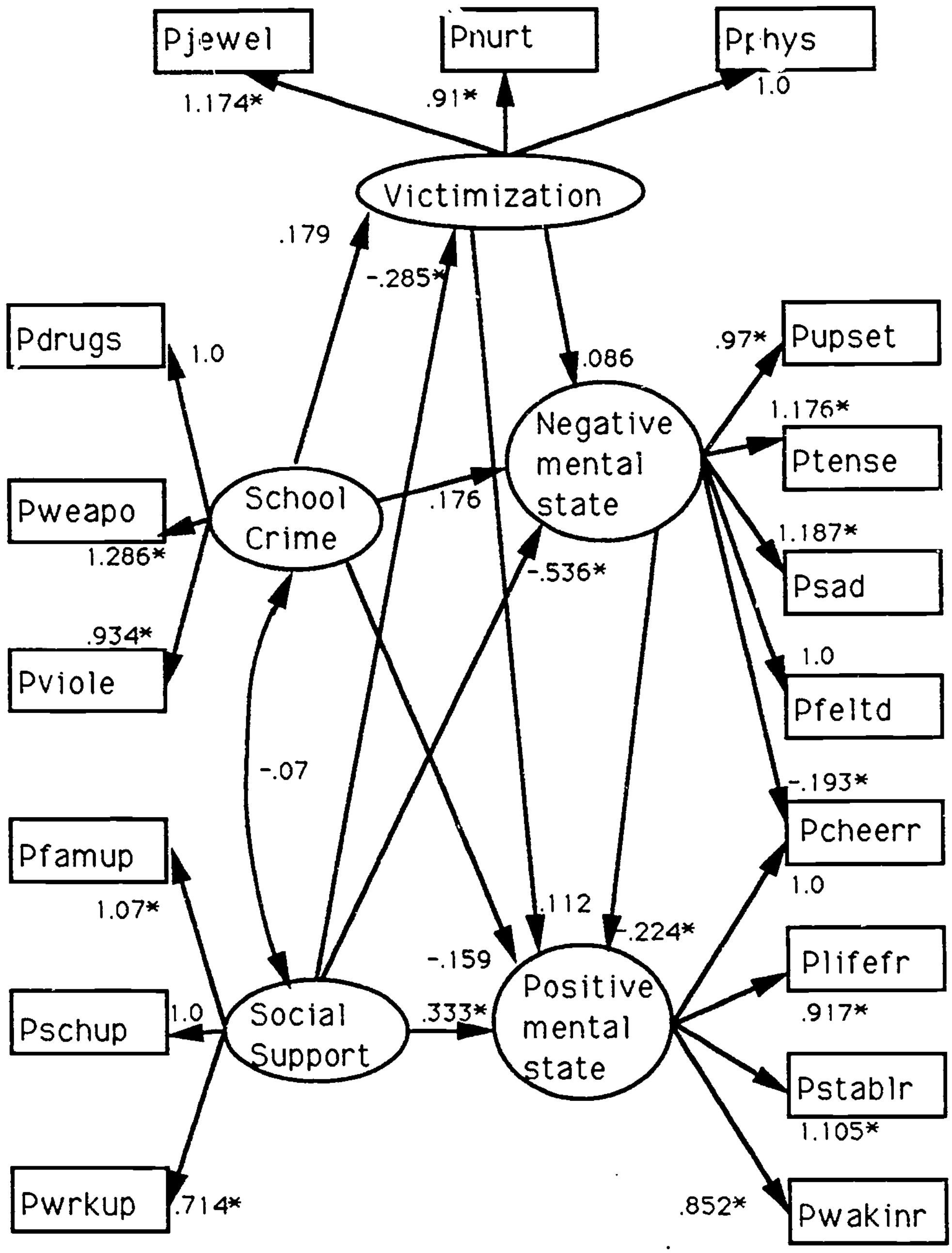
Partial Models

As we have stated previously, including pre-measured variables in our models reduced the effective sample size for the analysis. This reduction in sample size has decreased the power of our model tests, particularly for Model I; relatedly, the standard errors of parameter estimates are inflated. To examine the causal patterns estimated by the two general models with greater sample size, we used five partial models. Those models were analyzed with LISREL 6. The listwise deletion treatment for missin' values was employed for Submodels 1, 2, and 3; the pairwise deletion treatment was used for Submodels 4 and 5. The different treatment for missing data is because that Submodels 4 and 5 incorporated students' RCT data, which leave fewer cases matched in the sample. Submodels 1 and 2 had roughly a sample size of 200; Submodel 3 had approximately 151 cases, and Submodels 4 and 5 used the default value which was 100 in LISREL 6.

In order to compare roughly the estimates of the partial models to those of the general models, all the estimates of the partial models are unstandardized. Figure 10 shows the Submodel 2: social climate and mental health. There are three significant arrows starting from social support: social support to victimization (-.285), to negative mental state (-.536) and to positive mental state (.333). We may recall that th victimization variable has an effect on negative mental state in Figure 8. This effect disappeared in Figure 10. The difference between these two models may be due to three factors: (1) the incorporation of social support, (b) the omission of the variable "Conflict" and (c) different samples actually used in different models.

Figure 11 presents the Submodel 3: social climate and self-evaluation. We found that most results of this partial model are consistent with our general models, except for the arrow from victimization to self-esteem (-.159), which is absent from Model II. In checking Table 13 (Appendix B), we found the corresponding coefficient Beta (3,2) in Model II

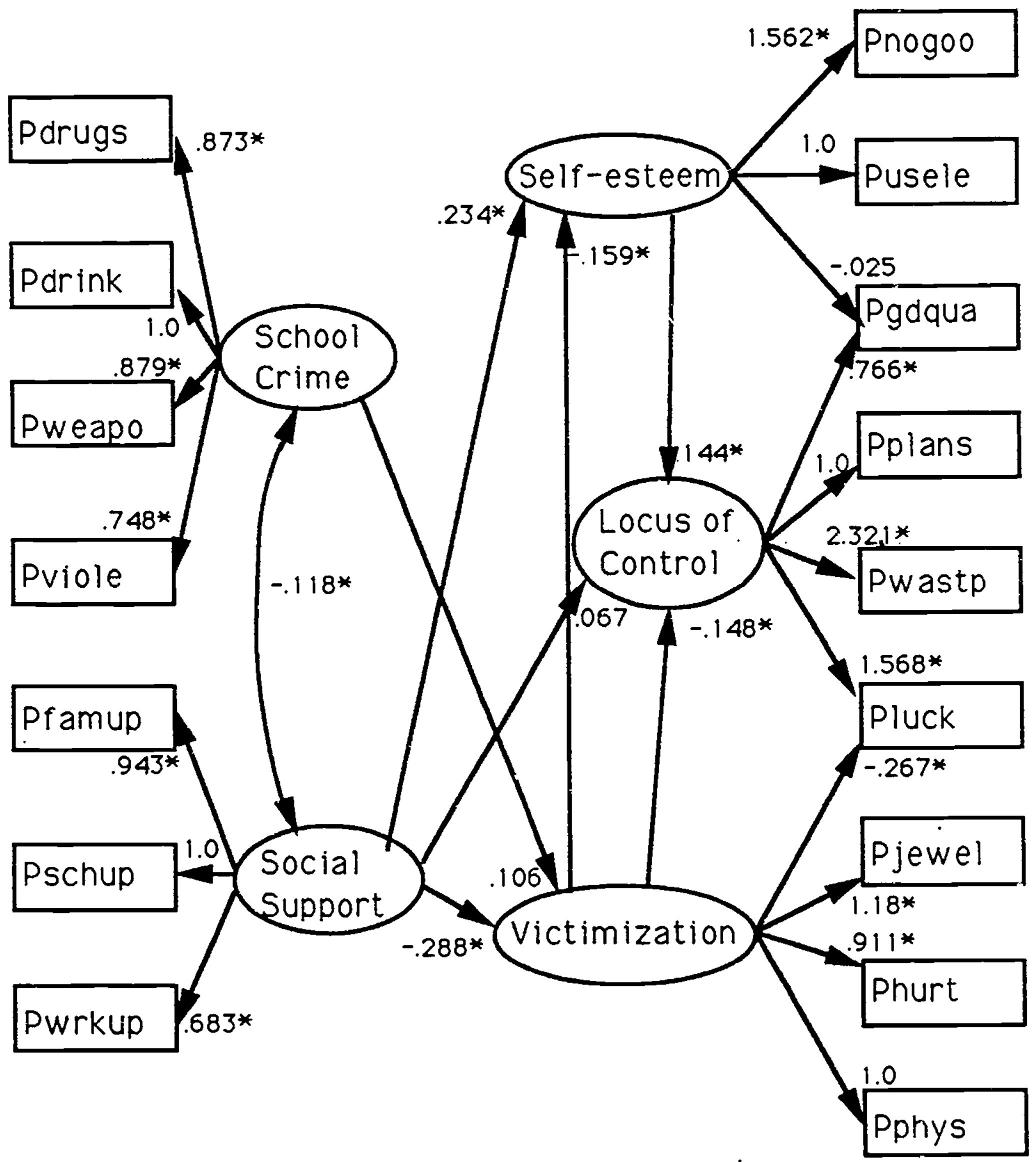
Figure 10
Submodel 2: Social climate VS. Mental health



Model fit: P = .207 for the MLE
Variance explained by the structural equations: .249
Goodness of fit index: .935
Adjusted goodness of fit index: .908
Root mean square residual:. 081

* t ≥ 1.96

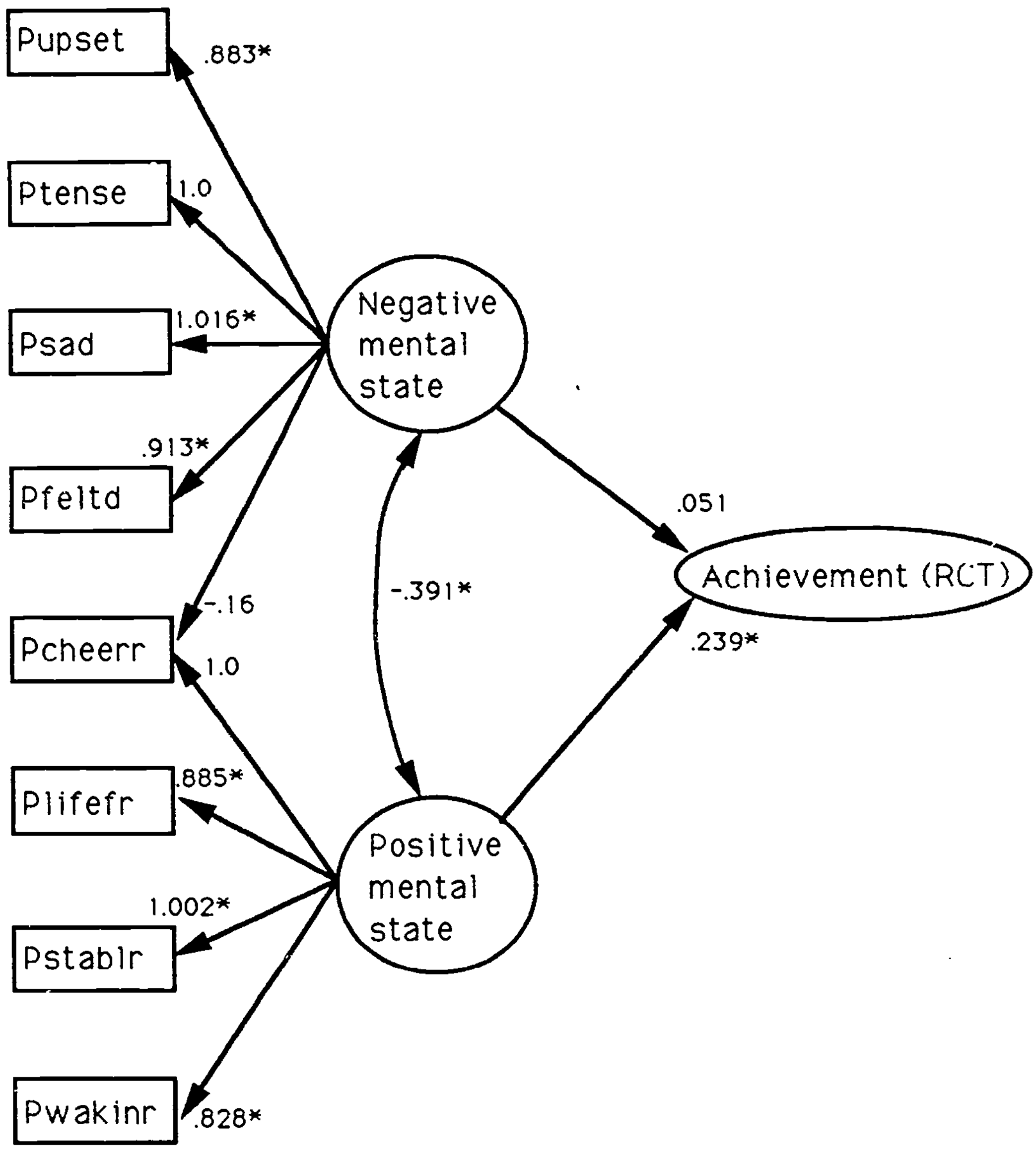
Figure 11
Submodel 3: Social climate VS. Self-evaluation



Model fit: P = .113 for the MLE
Variance explained by the structural equations: .178
Goodness of fit index: .936
Adjusted goodness of fit index: .907
Root mean square residual: .044

* t ≥ 1.96

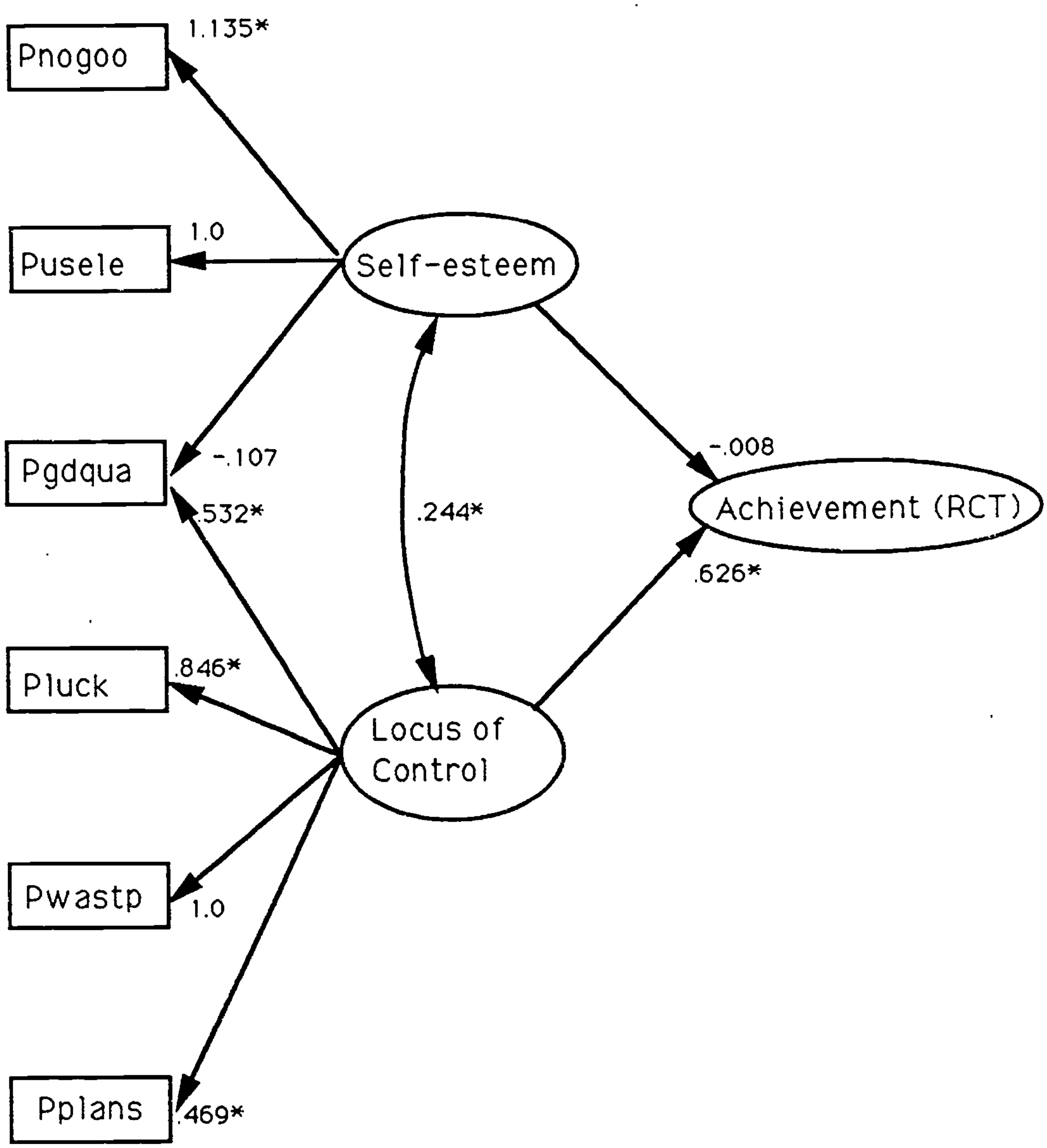
Figure 12
Submodel 4: Mental health VS. Achievement



Model fit: P = .977 for the MLE
Variance explained by the structural equations: .071
Goodness of fit index: .973
Adjusted goodness of fit index: .95
Root mean square residual: . 078

* t ≥ 1.96

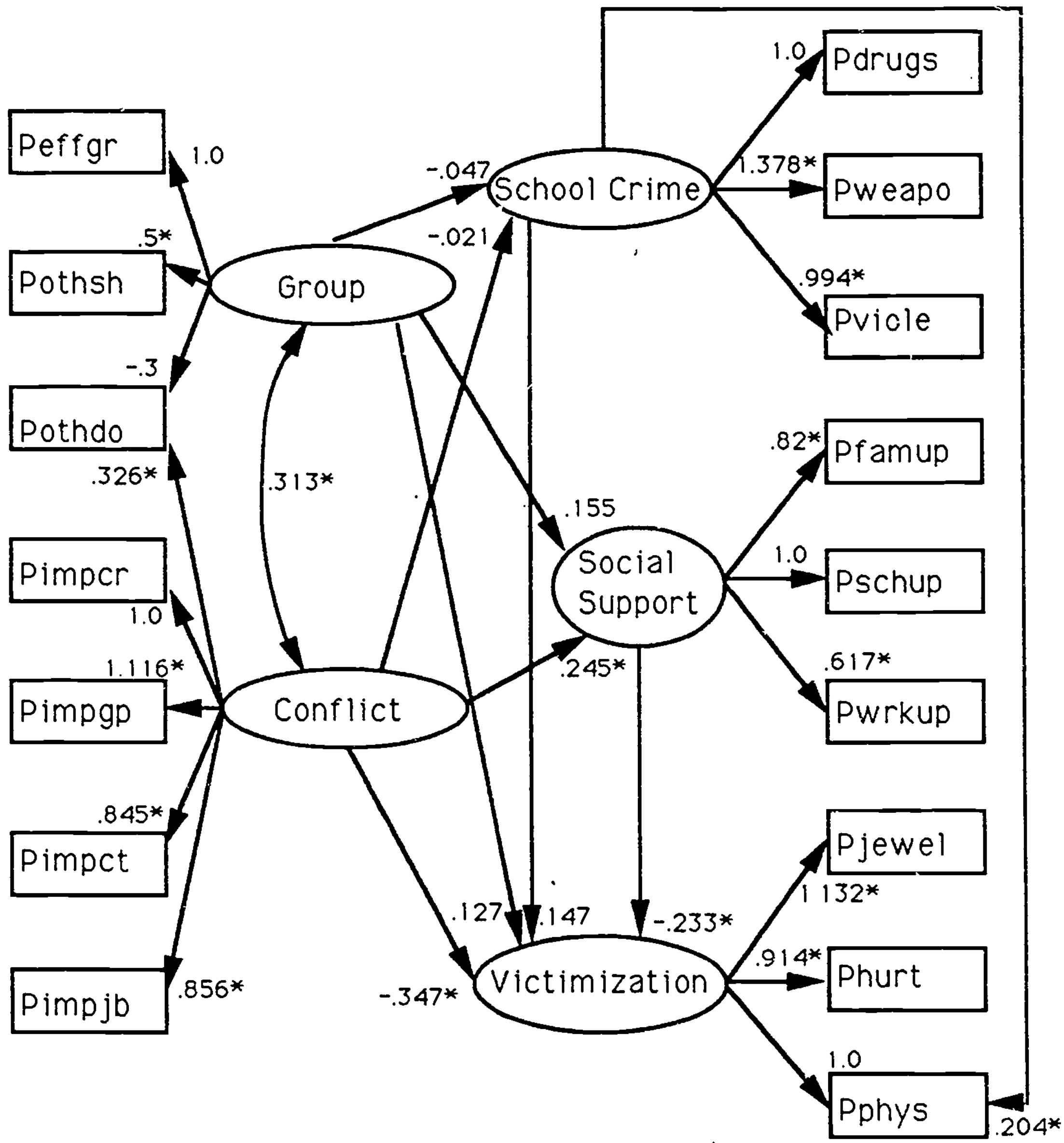
Figure 13
Submodel 5: Self-evaluation VS. Achievement



Model fit: P = .295 for the MLE
Variance explained by the structural equations: .212
Goodness of fit index: .964
Adjusted goodness of fit index: .908
Root mean square residual: . 047

* t ≥ 1.96

Figure 14
Submodel 1: Intervention VS. Social climate



Model fit: P = .102 for the MLE
Variance explained by the structural equations: .227
Goodness of fit index: .920
Adjusted goodness of fit index: .882
Root mean square residual: .056

* t ≥ 1.96

was -.089 with a standard error (.106), one half of the size of the one in Submodel 3. The reduction from Submodel 3 to Model II seems more likely due to the differences in specification rather than simply the difference in sample sizes or sample fluctuation. We see in Model II, that the variable of self-esteem has a unstandardized stability coefficient (Gamma (3,1)) of .375 (Table 13), and the variable "Conflict" has an effect of .239 on "Pesteem" (Gamma (3,5)). These two variables were not included in Submodel 3.

Submodel 4 in Figure 12 presents the effect of mental health on academic achievement. The effect of positive mental state on achievement was not found in Model I. Submodel 5 in Figure 13 provided substantial support for the effect in Model I of locus of control on academic achievement (.626).

Submodel 1 in Figure 14 supported the two general models for the effects of "Conflict" on the two intervening variables, social support and victimization. The effect of social support on victimization was not estimated in the general models, however, this effect is consistently estimated throughout the three Submodels 1, 2, and 3.

We may notice from Figures 10 to 14 that all the submodels fit the data well. The p-values of chi-square statistics range from .102 to .977, and GFI's range from .920 to .973. Except for Submodel 4, which is not confirmed by the general models, the rest of the partial models all have decent R-squares for the respective structural equations ranging from .178 to 249, given these are much smaller models compared to the general models.

In addition to verifying the general models, part of our underlying motivation for estimating these five partial models is to provide a series of flow diagrams, thus to test the hypothesized causal chain. If the interventions of conflict resolution and cooperative learning have yielded effects, their social psychological consequences should follow. And these consequences should be reflected in the students' academic achievement.

On the other hand, Deutsch's crude law also predicts a causal chain in the opposite direction. As we planned in our design, this prediction may be examined by checking the correlations between the pre-measured variables and the intervention variables, or the Gamma coefficients between the pre-measured variables and the climate variables in the general models. Table 10 indicates that there is a significant correlation estimate (.406) between previous locus of control and constructive conflict resolution (Phi (7,2)). Figure 5 shows that previous negative mental state may potentially reduce social support measured at a later time (-.26). Figure 8 suggests that the previously measured internal locus of control of a student may potentially reduce his/her victimization measured later (-.16). As mentioned above, the potential effect (-.16) of previous negative mental state on victimization measured later has a direction contrary to our understanding of the process. The PHI matrix of Model II (Table 15, Appendix B) reveals two significant correlations between previous locus of control and later improvements of conflict resolution (Phi (5,2)) (.402), and later improvements of effective group working (Phi (4,2)) (.396).

Model III

Our last two models were designed for detecting the intervention effects, which are often referred to as (Factor) Mean Structures. To estimate factor means, four more matrices have to be defined in addition to the earlier eight matrices. They are Tau-Y (τ_y), Tau-X (τ_x), Alpha (α), and Kappa (K). The general LISREL model is defined now by the following three equations:

$$\mathrm{ETA} = \alpha + \beta\mathrm{ETA} + \Gamma\mathrm{KSI} + \mathrm{PSI}$$

$$Y = \tau_y + \mathrm{Lambda}_y\mathrm{ETA} + \epsilon$$

$$X = \tau_x + \mathrm{Lambda}_x\mathrm{KSI} + \delta.$$

The α, τ_y, and τ_x are vectors of constant intercept terms. K, which does not appear in the equations, is a vector of means of KSI. The fit function to be minimized is now expressed as

$$F = \sum_{g=1}^{G} N_g F_g / N, \qquad \text{where}$$

$$F_g = .5(s^{(g)} - \sigma^{(g)})' W^{-1}_{(g)} (s^{(g)} - \sigma^{(g)}) + .5(\bar{z}^{(g)} - \mu^{(g)})' V^{-1}_{(g)} (\bar{z}^{(g)} - \mu^{(g)}) \qquad \text{where}$$

$$\mu^{(g)} = (\mu^{(g)}{}_y, \mu^{(g)}{}_x)' \text{ and } V_{(g)} = \hat{\Sigma}_{(g)} \text{ for ML.}$$

Where $\mu^{(g)}$ is a vector of the population means of all the manifest variables for group g, and $\bar{z}^{(g)}$ is a corresonding vector of sample means for group g. Thus, the overall fit funciton F is a linear combination of the separate fit funcitons for each group weighted by respective group sizes. Moreover, the F function for mean structures also take into consideration the differences between sample means and the estimated means, in addition to the differences in variances and covariances.

A fundamental difference between the analysis of a single group and the analysis of multiple groups is that in the analysis of a single group, the mean of a latent variable is undefined (not identified); while in the case of multiple group analysis, the means of a latent variable in different groups can be defined with respect to each other. Thus, in our situation, one group may be compared to another in terms of the effects of the interventions.

There are different ways of defining the factor means. As described previously, the origins and units of measurement of latent variables in a single group analysis can be easily defined by selecting a reference variable for each latent variable. In order to compare the means across groups, however, we need to fix an origin for each of the latent variables. Joreskog and Sorbom (1989) suggest that the mean of KSI in one group (such as the control group) be fixed at zero, and leave the rest free to be estimated. This will suffice, since all we may desire is the differences between groups. Similarly, the origin of α may be set in the same manner, when a model invloving ETA equations is concerned.

Bollen (1989) recommended that the scale for KSI, as well as Kappa, be assigned by setting Lambda (i,j) of i_{th} observed variable to one and the intercept (e.g., τ_{xi}) for the same variable to zero. Thus we have $E(X_i) = K_j$, where K_j is a single latent variable that underlies the observed X_i. The KSI_j is then given the same mean and units as X_i.

We followed Joreskog's convention in our analysis and set Kappa of Site 2 at zero. All the latent variables are treated as KSI's. As a result, there are no Alpha terms involved. Since the input matrices were produced by regular correlation program of SPSS, the pairwise missing information was not directly provided. Thus the number of observations for all sites were set equal to 70. (The sample sizes were later checked as 51 for Site 1, 56 for Site 2, and 73 for Site 3.) The results are presented in Figures 15-17.

Model III is a multi-group comparison of factor structures at the base-line level, that is, none of the Lambda X, Theta Delta, and PHI matrices are restricted by equality constraints, except where the factor patterns are fixed to be the same across groups. This model was designed to reveal the factor patterns at each site for the two intervention variables and the three intervening variables. It should be made clear that the parameter estimates among the latent variables in this model are only factor correlation coefficients, there is no causal implications attached to them.

Figure 16 displays the results for Site 2 which had both types of interventions. Compared with the other two sites (Figures 15 and 17), it is possible that the combined interventions at Site 2 has made the most systematic changes among these variables. Particularly, the correlation between "Group" and "Conflict" reached .75. This may an indication that these two training programs have produced mutual facilitation effect.

Figure 15 Factor structure: Site 1

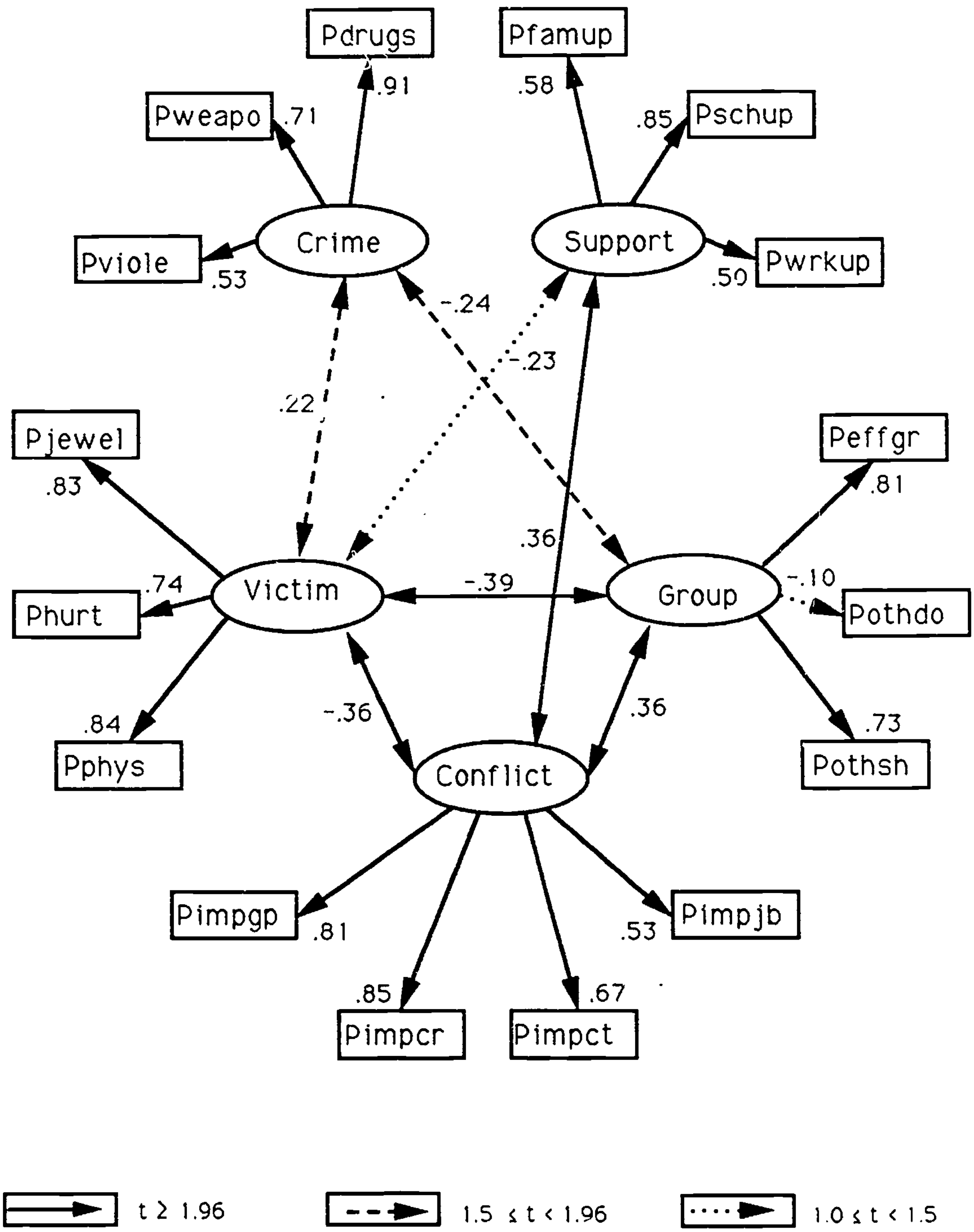


Figure 16 Factor structure: Site 2

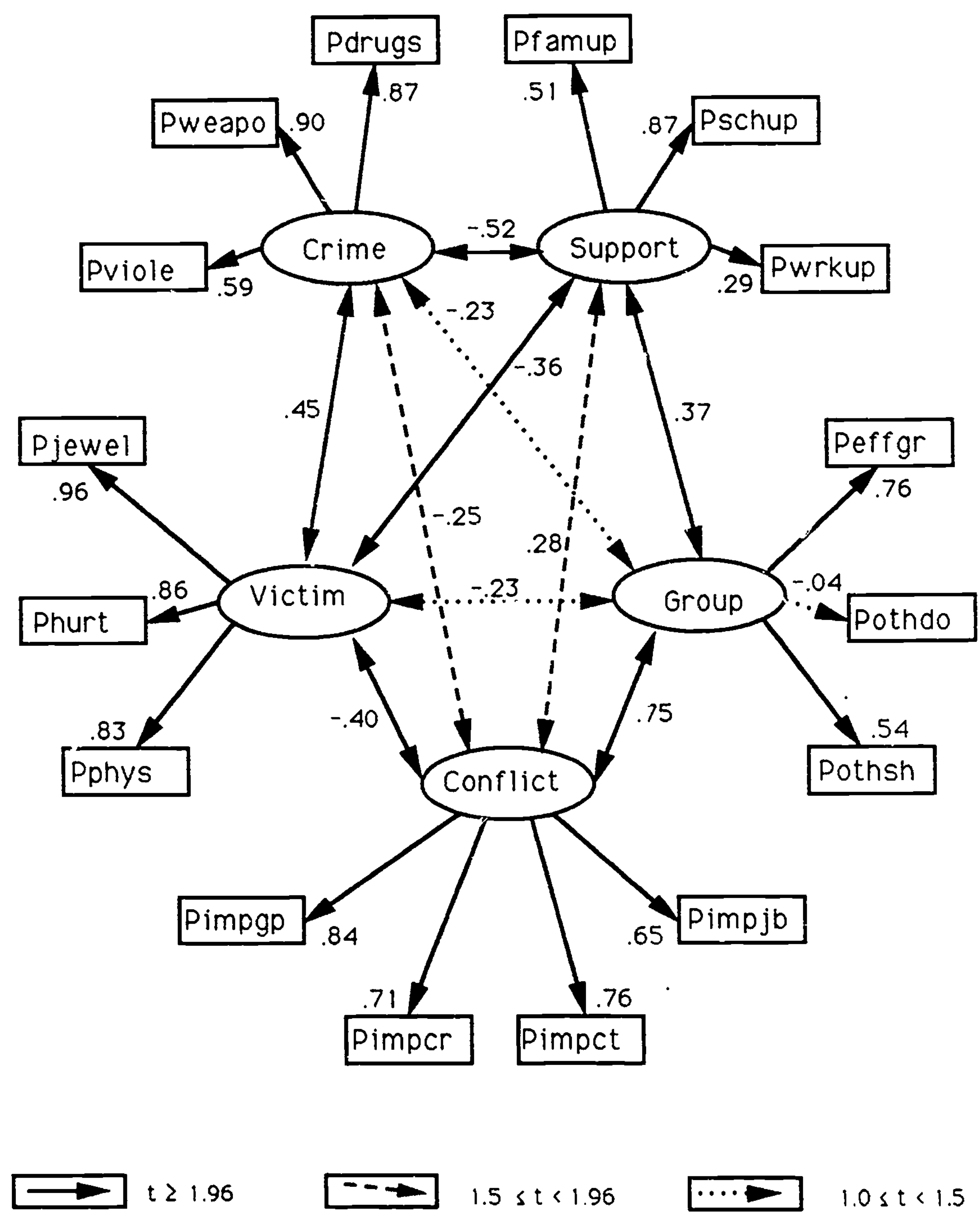
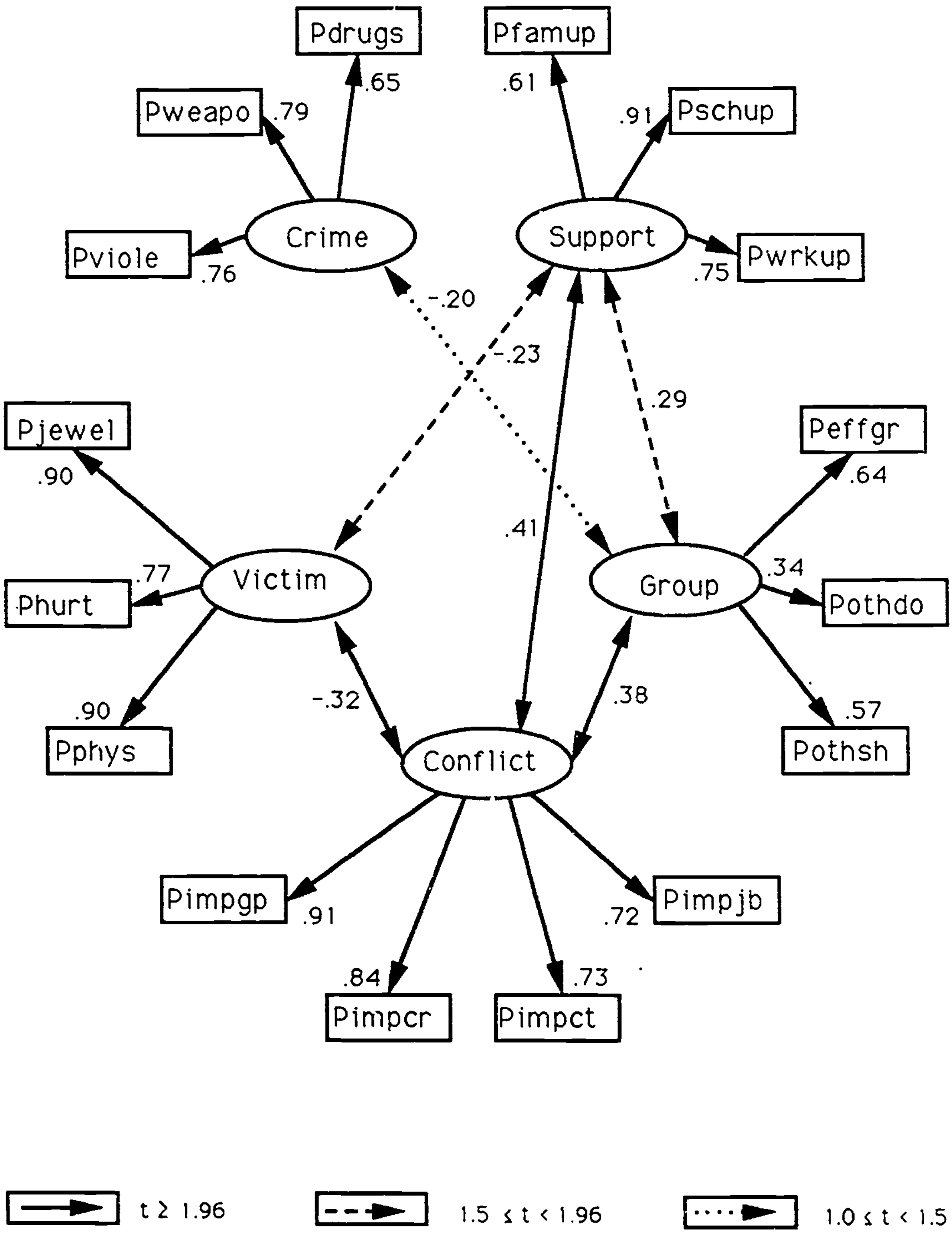


Figure 17 Factor structure: Site 3



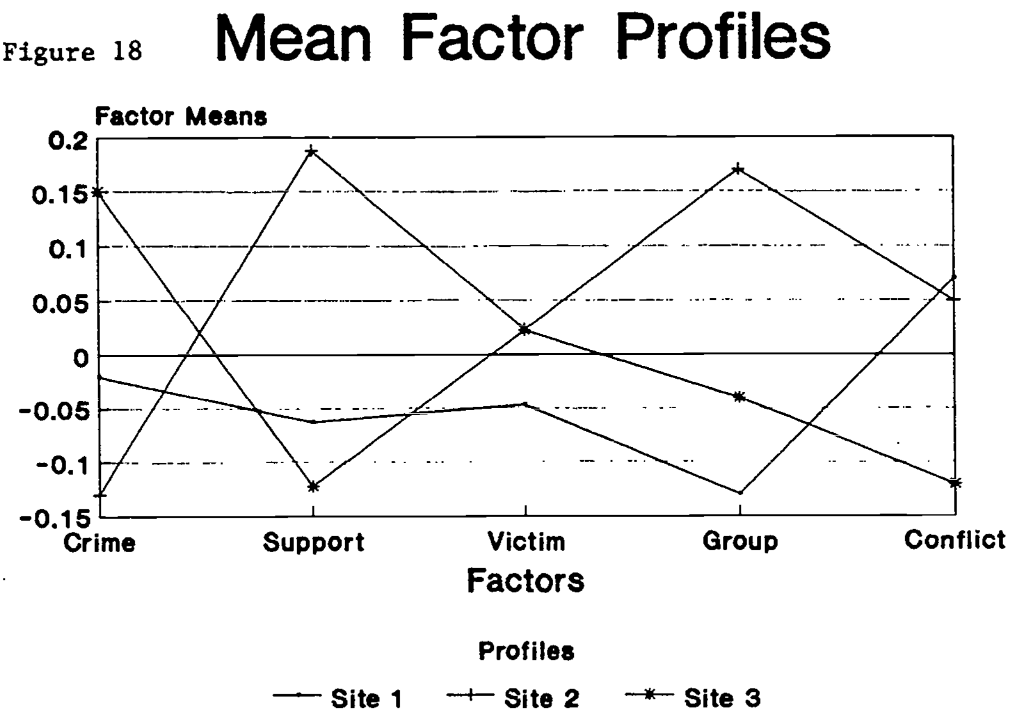
The overall chi-square for Model III with 305 degrees of freedom is 347.53, which corresponds to a t-value .047 (Table 20). This suggests that the model may not fit the data well. However, as we discussed before, the chi-square statistic is affected by a number of factors, and can only be treated as reference information. The constraints we set on Tau-X's to identify the model would also affect the model fit. Tau-X's were set invariant for the first two sites and zero for the third site, which were arbitrary constraints. There is no justification for the way these constraints were imposed in this model, we only intended to have the model be identified. As far as we understand, how to set constraints on the intercept terms should not affect the covariance structures among the latent variables. However, this may be wrong.

In general, more systematic variation is indicated among the variables for Site 2 than the other sites. However, to demonstrate that these differences are truely attributable to the interventions rather site differences, we need further evidence.

Model IV

Figure 18 reveals the factor mean structures of the five variables for our fourth model. This is an effect analysis model with which we evaluate the effects of the interventions. In this model, Lambda X, Theta Delta, and Tau-X were set invariant for all three sites. Since we

Figure 18


Mean Factor Profiles
Factor Means
0.2
0.15
0.1
0.05
0
-0.05
-0.1
-0.15
Crime
Support
Victim
Group
Conflict
Factors
Profiles
Site 1
Site 2
Site 3

intended to compare the mean differences, it would be reasonable to maintain measurement equality across sites. Kappa elements were set to zero for Site 2 so that a measurement origin for the factor means may be defined. Our emphasis is on the two intervention variables. It should be noted that "gain-scores" were used as the indicators of the intervention variables in all our models, thus possible pre-test differences have been taken into account. The methodological assumption of this particular analysis is that, if the factor mean structures could recover the original design pattern of the project from the data, then it would be reasonable evidence for the claim that the interventions of cooperative learning and/or conflict resolution had yielded effects. How strong the evidence could be depended on to what extent the two structures--the designed and the empirical--coinside.

Figure 18 presents the results based on scaled factor mean scores. The scaled mean scores were computed such that the weighted mean over the groups is zero for each factor. The method used follows an example by Joreskog and Sorbom (1989). Since we have set the sample sizes all equal to 70, the weight for each site then is a unity. Figure 18 shows that for the variable "Group", Site 2 has the highest mean score, Site 3 is in between, and Site 1, where cooperative learning was not introduced, has the lowest mean score. In contrast, for the variable "Conflict", Site 1 has the highest mean score, Site 2 follows, and Site 3, where conflict resolution was not implemented, has the lowest mean score.

Given the above mean structures, what is the probability that this pattern may be due to pure chance rather than true effects of the interventions? Since school sites were randomly assigned to each of the three treatment conditions, this probability may be calculated based on a randomization procedure. There were three positions for each intervention variable. The full permutation is 6 possible arrangements for each variable. With two variables, we then have 36 arrangements altogether. However, there were only two arrangements acceptable for each variable according to the original design, this is, either Site 3 takes the first position and Site 2 takes the second position, or Site 2 takes the first position and Site 3 the second on "Group", <u>and</u> Site 1 takes the first position with Site 2 the second, or Site 2 takes the first position and Site 1 the second on "Conflict". Thus the probability of obtaining an acceptable pattern is one out of nine times.

However, considering the qualitative analysis conducted independently by Mitchell (1991), we have found that the implementation of cooperative learning intervention at Site 3 has much lower quality compared to the same intervention implemented at Site 2. According to this result, the acceptable arrangements on "Group" may be narrowed down from two to one, that is, only the result with Site 2 standing at the first position and Site 3 the second will be acceptable for the variable "Group". Then the chance probability of getting the above obtained pattern may be reduced to one out of eighteen times, that is, .056. Certainly, this randomization test only provides a rough idea about the effect sizes, since the distances between positions in the diagram were not taken into consideration.

The implementation problems are, in fact, suggested in Figure 18. Since the cooperative learning intervention was implemented at both Site 2 and Site 3, their relative standing on the variable "Group" should be rather close, as is the case for "Conflict". The mean score of Site 3, however, falls far below Site 2. This result may be taken as a suggestion that the implementation of cooperative learning at Site 3 was not very successful, or the intervention of cooperative learning without conflict resolution was less effective.

The factor mean of Site 1 on "Group" has a t-value of -2.05. Since the factor mean value for Site 2 is fixed to zero, the t-value may be directly treated as the result of a two-group comparison. The result indicates that the factor mean of Site 1 is significantly lower than the factor mean of Site 2 which is zero. The t-value of the difference between Site 1 and Site 3 on "Conflict" is not readily obtainable from the LISREL output. Although they can be hand calculated by using the estimated factor variances and factor means for each site, we did not perform the computation. The overall result pattern which fits our original design well is perceived as sufficient information regarding this model.

The implementation problems are, in fact, suggested in Figure 18. Since the cooperative learning intervention was implemented at both Site 2 and Site 3, their relative standing on the variable "Group" should be rather close, as is the case for "Conflict". The mean score of Site 3, however, falls far below Site 2. This result may be taken as a suggestion that the implementation of cooperative learning at Site 3 was not very successful, or the intervention of cooperative learning without conflict resolution was less effective.

The factor mean of Site 1 on "Group" has a t-value of -2.05. Since the factor mean value for Site 2 is fixed to zero, the t-value may be directly treated as the result of a two-group comparison. The result indicates that the factor mean of Site 1 is significantly lower than the factor mean of Site 2 which is zero. The t-value of the difference between Site 1 and Site 3 on "Conflict" is not readily obtainable from the LISREL output. Although they can be hand calculated by using the estimated factor variances and factor means for each site, we did not perform the computation. The overall result pattern which fits our original design well is perceived as sufficient information regarding this model.

It was expected that the chi-square for Model IV would be lower than the one for Model III, because of the invariance constraints set for Lambda X, Theta Delta, and Tau-X across all three sites. The chi-square is 442.92 with 359 degrees of freedom (Table 24, Appendix B). The separate goodness of fit indices of each group for Model IV are all lower compared to Model III (Tables 18, 20, 22, and 24, Appendix B) . However, the fact that the results of the model conform to our research design is a more important criterion for judging the appropriateness of the model.

For the purpose of evaluation, we constructed two students' exposure measures. One was based on the proportion of time a teacher used cooperative learning in his or her classroom, and the percentage of time a student spent with that teacher; the second was based on the trainer's ranking of the teachers' effectiveness in implementing cooperative learning and/or conflict resolution, and again, the percentage of time a student spent with that teacher. However, we were not able to incorporate these two measures into our LISREL modeling, because the matched files did not have enough cases for running analysis.

Fortunately, Model IV provided sufficient and more reliable information than the exposure measures could have. Model IV has taken into account those factors which might affect the reliability of an exposure measure, and estimated the relative sizes of the effects of the interventions under the circumstances where the diffusion of the intervention effects, differences in quality and quantity in the application of the knowledge and techniques of cooperative learning and/or conflict resolution by different teachers might exist. Probabilistically, the findings of Model IV indicate the effects of the interventions over and above the influences of those factors.

Chapter Six

DISCUSSION

Methodological Implications

As we examined in the first two chapters, methodological controversies in testing Deutsch's theory of cooperation and conflict resolution, highlighted in the debate of Johnson et al. (1981, 1982), Cotton and Cook (1982) and McGlynn (1982), are in fact issues especially associated with traditional research methods. Due to the limited analytical capacity and the related limited research scope, the traditional approach of theory testing was unable to accommodate a large number of potential moderating/ mediating variables. Blalock (1979) has pointed out that it may be necessary to incorporate upwards of fifty variables in order to disentangle the effects of numerous exogenous and endogenous variables on the diversity of dependent variables that interest us.

Compared to previous research in this area, the present study is the first theory testing of its kind in terms of its conceptual scope, time span, and human and financial investment. The level of the test is the most general up to date, which includes testing the effects of three modes of training programs (cooperative learning alone, conflict resolution alone, and both together) and many of the major social psychological outcome variables concerning contemporary American adolescents have been accommodated.

There are several main differences between the traditional approach and the present confirmatory approach in this area. The first is the way of treating outcome variables. By the traditional approach, outcome variables are used only as dependent variables; thus their potential instrumental (intervening) functions may not be explicitly demonstrated unless simultaneous equations or path analysis is employed. In the present study, it is clear that the outcome variables such as self-esteem, locus of control, and academic achievement have functioned as both dependent and independent variables. In particular, self-esteem relays the effects of social support, which is affected by the conflict resolution intervention, to positive mental state and negative mental state in Model I, and possibly to locus of control in Model II. Academic achievement is affected by locus of control while itself in turn affecting self-esteem. Negative mental state are affected by social support and self-esteem, yet have detrimental impact on physical health.

With the advantage of estimating the intervention variables, intervening variables, and different types of outcome variables together in a non-recursive fashion, the identification of the coping resources for student's mental or physical health and the relationships between normally labelled as internal (self-esteem and locus of control) and external resources (social support) becomes feasible.

Secondly, rather than inferring causality from experimental control as in most previous studies, the causal hypotheses were directly tested regarding the effects of the interventions on the outcome variables. Previous studies normally utilized randomization to control possible confounding effects in experimental settings, or used statistical control for the effects of covariates in quasi-experimental settings (statistical control was also used in the true experimental settings in many cases); the methods of cross-lagged panel correlations (Campbell and Stanley, 1966), path analysis, or simultaneous equations were rarely employed in this area.

In the situation of randomization only, the exclusion of confounding factors is probabilistic. It does not always insure that the assumption of pre-experimental group equivalence holds, particularly when sample sizes are small or experimental cells are numerous. Due to the limited capacity of the traditional approach, statistical control is often inadequate in previous studies. ANOVA does not take care of the confounding issue. Dummy variable regression or ANCOVA does not accommodate multiple dependent variables. Normal multivariate techniques do not provide estimation of the relationships among the dependent variables. None of the traditional methods used in this area are able to model the structural relations among unobservable latent factors. However, without estimating the structural relations, it is impossible to reveal the true mechanism of how the students' mental health, physical health, self-evaluation, and academic achievement are affected by the interventions of cooperative learning and/or conflict resolution.

It is of course conceivable theoretically and mathematically that there is some possibility that a current estimated structure will be further modified by adding relevant new variables into the model. As we have stated, a scientific inquiry is an enterprise of endless approximation. Findings from one study are always ready to be verified or modified by other studies. The scientific community has long given up the Newtonian notion of absolute truth, since the prevalence or Einstein's challenge to the Newtonian cosmology. History has shown that no theory ever solves all the problems with which it is confronted at a given time; nor are the solutions already achieved perfect. As Kuhn (1970) perceived, "it is just the incompleteness and imperfection of the existing data-theory fit that, at any time, define many of the puzzles that charaterize normal science." (p. 146)

The important point of structural estimation is that it deals with questions, and thus provides answers, defined by a distinctive conceptual frame of reference. The causal assumptions are derived from previous knowledge; the estimations are confronted with alternative hypotheses as well as with empirical evidence.

For instance, as we have found that the conflict resolution intervention has a direct impact on self-esteem in Model II, while in Model I, such a direct effect does not exist; rather it is mediated through the variable "social support". This suggests that when the variable "social support" is incorporated into the model, the direct linkage between the intervention and self-esteem is explained away. Thus, it may be more accurate to attribute the changes in students' self-esteem to the enhanced social support resulted from their exposure to the intervention, rather than directly attribute those changes to the intervention itself. In other words, the level of the test of the Gamma parameter from "intervention of conflict resolution" to "self-esteem"--how many and what variables were incorporated in the entire model--provided a frame of reference for our judgment regarding the causal relations between the two variables.

Thus, the third methodological difference between the present study and previous studies lies in the fact that the existence of a conceptual and analytical frame of reference enables the present study to address issues concerning potential effects of mediating factors with greater certainty. The implication here may be seen by comparing Model I and II with the five partial models.

Examining Submodels 1, 2, and 3, we find that the effect of social support on victimization is consistently estimated across all the three models. Particularly, when the variable of conflict resolution intervention is included in the model (Submodel 1) with effects on both variables, the effect of social support on victimization remains significant. This gives us confidence to believe in the existence of that effect, although we did not estimate it in our two general models.

Comparing the general models with the partial models, we may claim that the general models are more reliable and more accurate. First of all, the general models included a larger number of variables than the partial models. Secondly, the general models used pre-measured variables as controls for the structural estimation. In other words, the general models have to a much greater extent taken possible confounding factors into account. Thus when an effect is estimated as significant in a partial model and consistent with the estimate in a general model, we then would consider with greater confidence the effect as being true.

Surprisingly, almost all the effects shown in the partial models were confirmed by the general models, except that the effect of victimization on self-esteem (Submodel 2 on Figure 11) and the effect of positive mental state on achievement (Submodel 4 on Figure 12) were explained away in the general models. Noticably, there is an effect of victimization, which is significant in the second general model but not in the submodel (Submodel 2 on Figure 10), on negative mental state. It seems that the inclusion of social support into the submodel bears the consequence, since the actual samples used in Model II and Submodel 2 do not differ very much. (Their effective sample sizes are 151 vs. 200, thus systematic missing data may not account for the difference.)

Theoretically, it is plausible to assume that whether an incident of victimization will cause depression or anxiety in a student depends upon whether he or she is able to reach out for social support. And people with more social support may suffer less from depression or anxiety and generally be more healthy than those with little social support regardless whether they have been victimized. However, there are some problems with this explanation. First of all, the most prevalent conceptualization regarding the effect of social support on mental health is the "buffering" theory (Brown and Harris, 1978). This theory may be represented by the following equation:

$$Z = \alpha + \beta X + \epsilon$$

$$\beta = \delta + \theta Y, \qquad \text{thus}$$

$$Z = \alpha + \delta X + \theta XY + \epsilon.$$

Where Z stands for mental health, X represents victimization, and Y is social support. δ is expected to be positive, and θ is negative, thus when social support is utterly lacking, the effect of victimization reachs its maximum. In other words, social support affects only the regression coefficient of victimization, and there is no main effect of social support on mental health in the equation. But it is indicated in Model I and Submodel 2 that there exists a significant main effect of social support on mental health either with or without the variable victimization. This implies a discrepancy between our finding and the more popular notion of "buffering".

However, I personally doubt that without a main effect, the interaction effect can exist in this case.

Secondly, Submodel 2 suggests also a common-cause structure among those three variables. A student's victimization and mental health may both be directly affected by the level of social support available to him/her; there is no main effect of victimization on mental health. This model has not been suggested in any other studies. However, given the special population we collected data from, the common-cause model may be plausible. This population consists of students mostly from black-African and Black-Hispanic backgound. They belong to the disadvantageous socio-economic classes. The school management thus has been oriented to promote social harmony and provide protection and support to its students rather than encourage competition among individuals. Consequently, the importance of social support is more salient and more emphasized among the students. The rate of victimization incidents is low (see Table 1), and the severity of victimization is much lower than in the outside society.

To further investigate the relationships between those three variables and to reach a better understanding of the underlying structure, a structural modeling incorporating latent interactive terms is probably necessary.

Stability of causal hypothesis testing, which is associated with the issue of measurement errors, is another characteristic that distinguishes the present approach from the traditional approach. The traditional approach in this area of theory testing basically takes the error-in-equation perspective; namely, only the error term of a regression equation can be estimated in its solution. Since variables used in the structural equations are disattenuated, that is, the variables in the causal structure are separated from their measurement errors, the structural parameters estimated are expected to be more stable across samples. This is called error-in-variable perspective. Although we did not directly test this methodological claim with independent samples, the consistency between the results of the general models and of the partial models suggests that the parameter estimates are fairly stable within this sampled population. In fact, the samples actually used for the general models are quite different from the samples for the partial models, due to the distribution of missing data on the variables involved.

However, the very advantage of the present approach with using disattenuated variables has also some drawbacks which deserve to be recognized. Bollen (1989) classified manifest variables into two groups: cause indicators and effect indicators. Cause indicators are observed variables that are assumed to cause a latent variable. For instance, in research on stress exposure, analysts often use measures of significant, disruptive events experienced by respondents, such as marriage, divorce, unemployment, and job promotions. According to Bollen (1989) these measures are cause indicators, since they cause changes in stress exposure. For effect indicators, the latent variable causes the observed variables. For instance, whether a person agrees with the statement "I feel that I am as good as the next person." is determined by his or her self-esteem, rather than the opposite.

Bentler (1980) viewed a latent variable caused by its indicator as a derived variable. He considered that this type of variables were not appropriate to causal modeling, since the uninteresting effects of random errors of measurement could not be separated statistically from substantively important causal influences in the model (1980). Burt (1976) indicated that the latent variables associated with effect indicators might suffer from confounded meaning which were different from what they were referred to in the structural modeling. Burt (1976) warned that "interpretational confounding occurs when the sources of empirical meaning used to interpret an unobserved variable are different from those used to estimate parameters in terms of which the unobserved variable is interpreted (p. 250)."

Burt (1976) identified five sources of empirical meaning assigned to an unobserved variable "D": (a) the indicator variables of the concepts which cause variation in D, (b) the indicator variables of the concepts which covary with D, (c) the indicator variables of the concepts whose variation is caused by variation in D, (d) the indicator variables of the concepts which covary with those concepts whose variation causes, covaries with, or is caused by variation in D, and (e) the indicator variables of the unobserved variable D itself. To varing degrees, an unobserved variable in a structural equation model is assigned empirical meaning by all indicator variables in the model.

An example in our LISREL modeling is the two intervention variables. The latent variables "Group" and "Conflict" are students' actual improvements in effectively working in groups and in handling conflicts with friends, peers, co-workers, and/or family members, which are supposed to be caused by the students' exposure to the interventions. Since the two constructs "effective group working" and "constructive conflict resolution" are closely related conceptually and operationally, the confounding in the meaning of the two latent variables are inevitable. The confounding is especially reflected in the factor structure analysis for Site 2 on Figure 16 where there is a correlation coefficient of .75 between "Group" and "Conflict". Thus it is very important to take the contribution from the students' exposure to the cooperative learning training or knowledge into consideration, when we examine the effects of "Conflict". Certainly, it may be practically impossible or even not necessary to disentangle all the potential confounding meaning contributing to the variable "Conflict". All that is important is to recognize that the structural modeling is a proximation of the unobservable process in which we have a research interest. The more we are aware of its limits and nature, the better our understanding is in terms of the true mechanism of that process.

The limitation of the structural modeling in theory testing manifested in our study may be viewed in the following aspects, in addition to the above issue of confounded meaning:

(a) Identification issue Perhaps the number one problem of dealing with a large-size model is the issue of identification. If an identification problem exists, even a correctly specified model may not be estimable. There is a necessary condition for the identification of all parameters:

$$t \leq .5(p + q)(p + q + 1), \quad \text{where}$$

t is the number of free parameters to be estimated in the model, p is the number of X-variables, q the number of Y-variables, and .5(p + q)(p + q + 1) is the number of independent elements in the observed covariance matrix.

However, the necessary condition does not provide sufficient guidance for either mathematical or empirical identification. Research has been directed to discovering sufficient rules for handling identification in structural modeling. Unfortunately, the sufficient conditions that have been proposed are almost impossible to verify in practice (Joreskog and Sorbom, 1989; Bollen, 1989).

Rothenberg (1971) showed that unknown parameters were locally (relative to global identification) identifiable if the information matrix is non-singular. This method of checking identifiability has been incorporated into the LISREL program. As a result, if a parameter is not identifiable, the analyst will be informed by the program in most cases. But, this check normally does not locate the sources of the problem. Thus, in the situation where the structural modeling involves large number of variables and factors, identification tends to be largely a blind procedure, and the identifiability of the model is not entirely under the control of the analyst.

Rindskopf (1983, 1984) analyzed a number of conditions based on the notion of empirical underidentification. A model that is algebraically identifiable does not guarantee that the model can be empirically identified. The causes of empirical underidentification can be numerous, such as nonnormality, nonlinearity, omission of an important path, a factor correlation close to one, or a factor loading close to zero, etc.

For instance, one of the rules for empirical underidentification is that a factor loading (either in the population or the sample) which is close to zero may cause instability of some parameter estimates by increasing their standard errors. With the large standard errors, those parameter estimates will result in wide confidence intervals. A likely consequence is that, probabilistically, the error variance estimates for those parameters may go beyond the legitimate boundary and become negative. This is called a Heywood case; for this phenomenon was first pointed out by Heywood (1931). In our analysis, the factor loading of "Pothdo" on Figure 5 is a suspicious case of this kind, the coefficient estimate of which is -.09. Although it is not algebraically proved, it is possible that this small factor loading has caused the instability of the error variance estimate of "Peffgr". The initial estimation of our models tends to have a negative estimate of the error variance for "Peffgr", but the estimate is positve for the maximum likelihood method.

A Haywood case is not particularly associated to the variable which resulted in a negative error estimate, and thus need not be eliminated by deleting that variable. Rindskopf (1983) provided several methods of parameterization to prevent the occurrence of Haywood cases. McDonald (1985) simply suggested that every factor be defined by three and preferably four or more variables having large loadings on it.

Blalock (1979) has indicated that realistic models of naturally occurring social phenomena must be nonrecursive or contain highly specific assumptions about lag periods or distributed lags. In other words, a realistic modeling of a social process should not only be structured in a nonrecursive fashion, but also contain one or more waves of data if possible. However, when a model is large and the algebraic procedure for identification is practically impossible, it is very difficult to have a clear idea about which of the reciprocal relations can be estimated mathematically and empirically.

When multiple waves of data are used, the problem of correlated error terms becomes salient, for the variables measured at different times tend to have correlated error terms. If variables measured at different times are not all specified as endogenous variables in the model, the potential correlated error terms may not be estimable; since Theta Delta and Theta Epsilon are separate matrices in LISREL. Thus the independence assumption of error terms is violated. The explicit consequence of this violation is a reduced model fit.

(b) Model evaluation issue The model evaluation has been a main problem in our study. As we have stated in the last chapter the model evaluation essentially has been affected by three factors: missing data, constraints on potential paths, and constraints on potential covariance between error terms.

Missing data There is no compelling reason in this study to suspect that the parameter estimates are systematically affected by missing data. Systematic missing data which have potential impact on the results may occur in two ways: (a) missing data associated with the intervention; (b) missing data associated with the outcome variables. In the first case, the missing data may reflect a differential attrition; that is, certain subjects drop out in the course of the intervention due to reasons which are related to the content, the methods, and/or the way it is implemented. In the second case, the sources of missing data may be selective subject groups; for example, students who have low attendence rate are more likely to be missing on the post-test. Due to their attendence rate, those students tend to have low academic achievement and thus low self-esteem.

The second type of missing data does not pose significant impact on the validity of our results. Since the study is mainly focused on causal modeling the effects of the interventions on students' self-evaluation, mental and physical health, and academic achievement, rather than on comparing the means of the outcome variables. One of the potential effects of this type of missing data may be its influence on the range of the scores of the outcome variables, which may suppress the strength of the association among the variables in the models. An examination comparing the group which took both the pre- and post-tests and the group which took the pre-test only showed virtually no differences on all the variables used in our models even almost without controlling for multiple comparisons.

The first type of missing data is less likely to occur in our study. The intervention uses an indirect exposure approach, and the students are exposed to the training and knowledge of cooperative learning and conflict resolution only through their teachers in their regular school classes. The content of the interventions is built into school curriculum.

Of course, besides the cases completely missing from the post-test among those who had taken the pre-test, a substantial portion of missing data occurred among the cases who took both the pre- and the post-test on different variables. This may potentially cause selective response patterns, which could lead to biased parameter estimates. However, the stability and consistency of the estimates across models may help us rule out this type of contamination in the results. This is another advantage of using partial models in the analysis.

But, the missing data have reduced the effective sample size, which may lead to the violation of the basic statistical assumptions. In general, LISREL estimation is based on large sample statistics. The statistical assumptions of ML estimators i.e. normality, and thus efficiency are only demonstrated as asymptotic model properties. If a sample size is too small, the fit function may not behave as chi-square distributed, thus the X^2 statistic may be incorrect. If the normality assumption is violated, the standard error estimates for each parameter may also be erroneous. Leamer (Bentler, 1980) suggested that the significance level chosen to evaluate a model be taken as a decreasing function of sample size, even if all the assumptions are met. However, there is no standard methodology existing for such choices so far.

Constraints on potential paths A constrained path in the structural equations explicitly indicates that there is no relationship between the two variables. As we stated before, for the sake of identification, some reciprocal paths are pre-fixed to zero, becasue those relationships are not of research interest.

The possible consequences concerning the validity of the results of a model with arbitrarily constrained paths depend on the status of the involved variables related to those paths. If the variables, on which a path is constrained, are hypothesized as having effects or mediating effects on other variables, then the accuracy of the parameter estimates related to those variables may be questionable. In the present study, the main constrained paths are the Beta coefficients between negative mental state and positive mental state, and these two variables are basically treated in our hypothesized model as outcomes. However, the path parameter from negative mental state to physical health was released for estimation based on empirical information (modification indices). Given the magnitude of the estimated effect (-.72), the existence of that path is not in doubt. However, the constraints on the paths between negative mental state and positive mental state are expected to reduce the goodness of fit for the models in which these two variables are both used.

Constraints on potentially correlated error terms As we indicated previously that a full LISREL model is defined by three equations. In order to solve the three equations, certain assumptions have to be made. These assumptions are:

(1) The disturbance terms in the structural equations are uncorrelated with any of the latent exogenous variables.

(2) The error terms for Y-variables are uncorrelated with their causes--the latent endogenous variables.

(3) The error terms for X-variables are uncorrelated with their causes--the latent exogenous variables.

(4) The disturbance and the two types of error terms are mutually uncorrelated.

(5) The matrix $I - \beta$ is non-singular.

Without the first four assumptions, the model would have more unknowns than equations and the model then could not be identified. However, if carefully specified, some of the above assumptions may be relaxed, and the covariances between some of the error terms may be estimated. EQS provides the option of estimating correlated error terms across matrices. Unfortunately, LISREL programing does not support this option.

In the present study, the pre-measured data are used as exogenous variables and the corresponding post-measured data are used as the endogenous. The possible correlation between the error terms of the pre-post data thus cannot be estimated by LISREL, unless some major change in the model specification is considered.

Given the existence of a large portion of missing data and those constraints in this study, the goodness of fit indices are used only as reference information.

Model evaluation problem is related to a more general problem, namely, how to control type I error in LISREL modeling. Fornell and Larker (1982) raise the issue that the t statistic, which tests each of the parameters in a model, may understate the overall type I error, since there are multiple parameters being tested in the model. However, what procedures should be used to control the overall type I error is not clear. Fornell and Larcker (1982) suggested that multiple comparison procedures be used, which would be Bonferroni adjustment in this case. However, Bonferroni adjustment is only appropriate when (1) the multiple dependent variables used are highly intercorrelated; (2) the comparisons are made at the same level. In other words, If two dependent variables are mediated or affected by different variables, then the Bonferroni control for the overall alpha is inappropriate for the two dependent variables. Using Bonferroni adjustment, in my opinion, would be too stringent in LISREL modeling.

As Joreskog (1982) has pointed out the purpose of LISREL modeling is to test the validity and adequacy of a model, which is based on a substantive theory or known results from previous studies, and to estimate its parameters. The statistical problem in LISREL modeling is to extract as much information as possible out of a sample of given size without going so far that the results are affected to a large extent by "noise". In general, it is the difference between X^2 that matters rather than the X^2 values themselves (Joreskog, 1982). This is so because in most practice, we are not sure whether all the basic assumptions are fulfilled or whether the sample size we have is appropriate.

There are a number of other issues in our study which may deserve further research. As discussed above, there are different opinions regarding cause indicators. To avoid more complication, we did not employ alternative modeling in which variables such as social support and victimization might be treated as latent factors determined by their cause indicators. In our present models, these two factors are taken as causes of their respective indicators. The justification is that the true social support a person has experienced affects his or her perception and memory of whether certain agents are being supportive; similarly, the true experience of being victimized determines his or her responses to the questions. But, if this is the case, then probably there will not be any cause indicators unless the data collected are from some reliable sources other than a respondant's memory. Bollen (1989) has suggested that the Lambda coefficient of a cause indicator be set to one, and the corresponding error term be zero, which implies that a cause indicator is error free. However, this way of specification often makes a model underidentified, since it tends to reduce the number of equations. Thus it requires further research to test the alternative models in which social support and victimization are treated as being determined by their cause indicators. The results may be compared with the present findings so that a better modeling may be identified.

Attention in further studies should also be led to different transformation methods which may be applicable to the indicator variables of victimization. Bollen (1989) empirically demonstrated that the logarithmic transformation tend to correct the extreme skewness and kurtosis. It may be worthwhile to see whether using transformed variables in the models would increase the explanatory power of these models.

Theoretical Conclusions and Discussion

The main theoretical findings in this study are reflected in Models I & II. The partial models provide corroborative evidence and supplementary information. Model III portrays the factor structures across the three experimental school sites, which are very informative regarding the differences in the implementation of the interventions. Model IV presents the profiles of factor mean structures for the effect evaluation against the original intervention design.

Two general issues in the findings of this study

One of the general findings of the entire analysis is that the study has demonstrated that most of the effects of constructive conflict resolution do not directly impinge upon a student's self-esteem, locus of control, mental health, physical health, or academic achievement. Rather, those effects are mediated by the changes in his/her interpersonal relationships indicated by a student's social support and victimization. Through his/her enhanced social support or reduced victimization, constructive conflict resolution may, on one hand, increase a student's self-esteem, locus of control, and positive mental state (i.e., his/her cheerfulness, general interests in life, emotional stability, and general activeness), and decrease his/her negative mental state (i.e., depression, anxiety, and sadness), on the other hand.

Social support is relatively a new concept in the literature of conflict resolution. The earlier research within Deutsch's paradigm on conflict resolution was mostly conducted at the micro-level and focused on motivational or attitudinal factors such as trust and suspicion, or factors which might have influence on the process of conflict resolution such as communication, etc. There was little research on the mechanism that transmitted the effects of conflict resolution onto its social psychological outcomes. Although it is, as we have stated before, a logical derivation of Deutsch's conflict theory, the moderating/mediating function of social support had not been brought under scrutiny in Deutsch's laboratory until late seventies (Steil, Tuchman, and Deutsch, 1978; Steil, 1980;

1983). In their study of the responses to injustice by advantaged and disadvantaged subjects, social support was found to lead the disadvantaged, but not the advantaged, subjects to become more sensitive to the injustice. Moreover, social support reduced the tendency for self-blame among the disadvantaged subjects. In other words, social support moderated the effects of conflicts which were manipulated as justice/injustice conditions in their studies.

Along with the increasing awareness of the importance of the application of psychological conflict theories to resolving conflicts at schools, conflicts at work places, union-management conflicts, urban interracial conflicts, conflicts in international trade negotiation, and international political conflicts, theories and strateges of conflict resolution have been widely studied. Thus the moderating/mediating function of social support emerged as one of the main research topics in the field of conflict resolution, since social support was found to be a powerful explanatory variable by many researchers of organizational behavior (Abdel-Halim, 1982; Jackson, 1983; Seers et al. 1983; Berkowitz and Gibbs, 1985; Quick, Horn, and Quick, 1986; Baer et al. 1987; Marcelissen et al. 1988; Jason et al. 1988; Kashani and Shepperd, 1990).

Concurrent with the experimental research carried out in Deustch's social psychological laboratory, early studies of the effects of conflict in conjunction with social support were being conducted by occupational/organizational oriented psychologists (Gavin, 1977; La Rocco, 1978; Warr, 1982; Strumpfer, 1981). However, except for the studies by Steil (1980, 1983), few of the earlier studies or later ones explicitly addressed the moderating or mediating function of social support between conflicts and their outcomes. The main reason for that is that the traditional analytical methods could not easily accommodate large number of dependent, independent, and intermediate variables together. Thus it was difficult to demonstrate the direct-indirect effects and all the relationships among those variables.

In a more recent study of role stress in business settings in which LISREL analysis was used, Schaubroeck, Cotton, and Jennings (1989) found that the effects of participation in decision-making and co-worker social support on job satisfaction may account for the direct effects of role conflict and role ambiguity on satisfaction and turnover intention observed in previous studies. Specifically, when social support and participation were brought into the model which included variables such as role conflict, job satisfaction, turnover intention, the direct effects of role conflict on satisfaction and turnover intention disappeared. Instead, social support and participation yielded direct effects on job satisfaction, and the latter contributed to the turnover intention. These findings support our results.

In the present study, when social support was absent in

Model II, the variable "Conflict" had direct effects on both self-esteem and negative mental state. However, when social support was incorporated into the model (Model I), all the effects of conflict resolution were mediated through social support. These findings suggest that it may not be the direct influence of conflict resolution that affect a student's self-evaluation or mental health; rather, it is the repeated successes in constructive conflict resolution which improve his/her interpersonal relationships, and the latter leads to better social support. Thus the enhanced self-evaluation and mental health may be the consequences of better social support which results from successful conflict resolution, rather than being the direct results of constructive conflict resolution.

The other general issue pertains to the test of Deutsch's crude law of social relations. Our results have strongly indicated that a student's mental health, physical health, self-evaluation, and academic achievement are positively affected by constructive conflict resolution. As we have indicated previously, to test whether these effects may further facilitate constructive conflict resolution among students and thus further promote better interpersonal relationships, the second and third waves of data should be analyzed. Since our modeling was based on the first and second waves of data, naturally, the covariation between the first wave control variables and the intervention/ climate variables did not come out as strongly as otherwise could have been expected if the second and third waves of data had been used.

However, the purpose of using the first and second waves of data in our causal modeling is only roughly to simulate the mechanism underlying the reciprocal relationship predicted by Deutsch's crude law. There are two additional theoretical reasons which may explain why the underlying machenism cannot be estimated accurately in this case.

(a) The underlying mechanism of Deutsch's crude law is process-specific. That is, the characteristic processes or outcomes most likely function as reinforcers for further interpersonal changes, if the given interpersonal relationship has been a cause of those characteristic processes or outcomes. In other words, an individual may tend to engage in promoting a given type of interpersonal relationship if he or she has experienced social psychological or educational consequences resulting from that type of interpersonal relationship.

(b) Since this type of test regarding Deutsch's crude law has never been conducted in the past, the model specification is basically intuitive and exploratory. There is no research findings which may indicate which of the outcome variables may affect that given type of interpersonal relationship and how the effect will occur. Thus we have no explicit rationale for specifying one parameter for estimation rather than another. This is why we set the entire Gamma matrix free in Model II to explore empirical relationships which might exist. In addition, the PHI matrices in both Model I and II are subdiagonal matrices, thus the pairwise correlations between the firstwave variables and the intervention variables are fully estimated. Moreover, control for type I error is not applied, which, as a general methodological issue we have discussed in the previous section.

Simply speaking, (a) stresses that using the first wave data may not be an effective way of simulating the second part of Deutsch's crude law; (b) indicates that the estimation of this part is empirically-driven rather than confirmatory.

The significant correlations between initial locus of control and improvements in both effective group working and constructive conflict resolution suggest that the maintained or enhanced internal locus of control may potentially function as a catalyst for further improvements in those two areas. Initial locus of control is also found to be potentially effective (-.16) in reducing the risk of being victimized, measured at a later time.

Generally, one key point for a study of social intervention is that the study should investigate not only whether the intervention has yielded effects and produced anticipated social changes during the course of the intervention, but also whether the induced changes will last and even create a social environment in which further changes may take place. External initiation of social changes is intended to ignite the internal engine so that further social changes can be re-energized through the internally initiated programs within that institution or even self-initiated action and exercise of the given individual him- or herself. This is the essence of Deutsch's crude law, in terms of its application to a social intervention and social change. This is also the reason why we based our overall research design on Deutsch's crude law.

Our findings regarding testing the second part of Deutsch's crude law is inconclusive due to the limitation of information available. Although the reciprocal pattern is suggested in the data, a causal specification rather than a correlation estimation is needed. Thus future study should employ the second and third waves of data so that the outcome variables in the second wave may be specified as causes of the intervention variables from the third wave.

Conclusions of the hypothesis testing

There are altogether 23 hypotheses in our original research design. Among them, the three hypotheses related to school crime were mistakenly included in the second general model. In fact, these hypotheses may not be testable given the data we collected. For testing them, schools have to be used as sampling elements rather than individuals. However, we left the variable "Crime" in our models as a heuristic evidence to demonstrate the sensitivity of LISREL.

One of the frequent concerns regarding the validity of causal modeling is that, granted a causal relationship exists between two variables, how do we demonstrate that the estimated coefficient reflects a causal effect between the two variables rather than simply an association in perceptions. For instance, if a person is physically attacked at school, will it be more likely that the person perceives crime at the school being more severe than others who were never attacked? Assuming school crime does cause a student's victimization at school, the parameter estimate between school crime and a student's victimization should be still around zero. This is so because school crime is a constant for a given school, while a student's victimization varies from person to person. On the other hand, if the estimaiton reflects only the association in perceptions, then the parameter estimate should be relatively high. In other words, the power of LISREL is to separate true school crime from the differences in each individual's perception; thus school crime becames relatively a constant (which varies only randomly) and is isolated from the rest of the variables in each model. Therefore, the nearly isolated variable "Crime" in the second general model and the partial models, to some extent, speaks for the validity and reliability of our measures and the strength of the LISREL techniques.

Excluding those related to school crime and Hypothesis

21 which is not tested, the entire analysis has tested the 19 hypotheses included in the original design. In addition, two effects have been suggested by the empirical relations among the variables in Model I, which are Beta coefficients from achievement to positive mental state and from negative mental state to physical health. The Gamma matrix is fully estimated in Model II. The total estimated coefficients thus are many more than 19.

Out of the nineteen original hypotheses (Table 25), nine are exactly confirmed with cross-validation; one is exactly confirmed without cross-validation; two are confirmed with t-values less than 1.96 in Model I (the power of these two tests are .34 and .28 respectively). Hypothesis 20 predicted effects of victimization on both negative and positive mental states, the effect on the negative mental state is confirmed in Model II but not in Submodel 2. Hypothesis 22 is confirmed in Submodel 3 but not in Model II.

Table 25

Check List of the original Hypotheses Confirmed

Hypotheses	Confirmation	Cross-validation
1	Yes	Yes
2	No	No
3	N/A	N/A
4	N/A	N/A
5	Yes	Yes
6	Yes	No
7	Yes	Yes
8	No	No
9	Yes	Yes
10	Yes	Yes
11	No	No
12	No	No
13	Yes	Yes
14	Yes[a]	No
15	No	No
16	Yes[b]	No
17	Yes	Yes
18	Yes	Yes
19	N/A	N/A
20	Yes[c]	No
21	N/A	N/A
22	Yes[d]	No
23	Yes	Yes

[a&b]These two coefficients are confirmed with t-values less than 1.96 in Model I (the power of these two tests are .34 and .28 respectively).

[c]This hypothesis predicted effects of victimization on both negative and positive mental states; the effect on the negative mental state is confirmed in Model II but not in Submodel 2.

[d]This hypothesis is confirmed in Submodel 3 but not in Model II.

Among the unconfirmed hypotheses, the more important one is Hypothesis 2 which predicts the effect of cooperative learning on a student's social support. Deutsch (1949a, 1949b, 1958, 1960, 1962, 1973) has predicted that promotive interdependence produces better communication, mutual help, mutual trust, mutual commitment and sense of responsibility among cooperators. Johnson and Johnson (1989) in their survey of past research has concluded that cooperation generally will produce greater social support than competitive or individualistic efforts. It seems that improvements in effective group working should be expected to induce better social support.

There are possibly several reasons why this expected effect did not occur. First of all, the variable we used for measuring cooperative learning in our analysis is different from those used in previous studies. Most of the past studies did not separate the component of conflict resolution or interpersonal problem solving from their manipulation and/or measures of cooperative learning. The effects of cooperative learning and conflict resolution are likely to be intermingled with each other in those studies.

Certainly, it may not be possible or even conceptually plausible to separate the two variables entirely. Effective group working requires participants in a group to be in social harmony; constructive conflict resolution may be reflected in whether the students are able to work together cooperatively in the group. Thus, although these two variables are different theoretical constructs, they are closely related to each other. The problem is how we can accurately operationalize the distinctions and connections. By LISREL modeling, we are to some extent able to select the indicators which may better represent the entangled relationship and incorporate each component in the two latent factors with appropriate proportion.

Unfortunately, the latent factor "Group" is not effectively indicated by the manifest variables; the indicator "Pothdo", which is supposedly to measure self-expectation of help from others in a group, has equivocal meaning and is poorly worded. That left only two items indicating the factor "Group." Thus the variable of effective group working is not well operationalized in our models.

The implementation of cooperative learning intervention at Site 3 (which is the cooperative-learning-only site,) is generally not successful, which has been suggested in the results of Model IV. The qualitative analysis (Mitchell, 1991) of the intervention project indicates that the implementation problem may be attributed to several main factors: less training time at this site than the others; a trainer who is very experienced with the content and methods of cooperative learning but is inexperienced with schools in an urban setting where minority students consist of the main population; a relatively anthoritarian school management as compared with the other two school sites; and a high student dropout rate.

The students' exposure to the intervention at Site 3 is much less systematic and consistent compared to the other two sites, which is indicated by its lowest three-consecutive-cycle retention rate. The retention rate for cycle 4 at Site 3 is .41 as compared with .34 and .53 at the other two sites; the retention rate for cycle 5 at Site 3 is .21 as compared with .59 and .69 at the other sites; the retention rate for cycle 6 is .22 as compared with .59 and .68 at the other sites. The low attendence at Site 3 would surely impair the function of cooperative learning intervention.

Effective group working is based on the capability of group participants of managing problems or conflicts constructively. Cooperative learning training alone may be less effective than cooperative learning combined with conflict resolution as is the case at Site 2. If this is correct, even if the cooperative learning training has produced effects at Site 2, which is actually demonstrated, the overall results may still not be significant since the training at Site 3 was not effective and Site 1 only received training in conflict resolution. Clearly, if cooperative learning training did not create systematic changes for the overall sample, the impact of effective group working on social support would not show a systematic pattern.

Model II shows a possible effect of effective group working on negative mental state with a magnitude of .11 which is the smallest among all the causal estimates. It may be due to the fact that when rules of cooperative learning are not well implemented or understood, or when the goal interdependence in a cooperative learning setting is not fully realized and appreciated by the participants, cooperative learning may produce tension and anxiety. The better-achieving students may develop fear of being hindered or penalized by their fellow students in a cooperative learning situation; while the less-achieving students may feel pressured by being responsible for the group performance when the cooperative learning rules are introduced.

Effective group working is seen as possibly increasing the internal locus of control among students (Model II). It may be conceivable that increased social responsibility and responsibility for the achievement of other group members in the cooperative learning situation enhances a student's sense of self-efficacy and responsibility for him- or herself. Symbolic interaction theory indicates that intrapersonal processes echo interpersonal processes (Stryker, 1985); thus the congruence of the group members' social responsibility and their responsibility for themselves should be reflected in their increased internal control.

In the methods chapter, we have presented the underlying philosophy of the analytical design based on

Deutsch's crude law. We may now integrate the empirical results following the overall design and present them as follows:

(a) the effects of constructive conflict resolution on social support and victimization have been demonstrated in Model I and II, and supported by Submodel 1. Conflict resolution may also have direct impact on the student's mental health and self-esteem (Model II).

(b) The effects of effective group working on social support and victimization are not directly confirmed. However, the correlation between the two variables "Group" and "Conflict" suggests that there is certain amount of contribution from effective group working to the effects of constructive conflict resolution on other variables. Furthermore, effective group working has been suggested to have an impact upon some outcome variables (Model II). But the operational, implementational, and other problems indicated above prevent us from making definitive conclusions in this regard.

(c) Changes in interpersonal relationship, elicited by constructive conflict resolution and reflected in the measures of social support and victimization, are indicated to be effective in promoting mental and physical health, facilitating positive personal attitudes toward life, and enhancing self-esteem and internal locus of control (Models I and II, and Submodels 2 and 3).

(d) Constructive conflict resolution, mediated by school social climate and locus of control, enhances a student's school academic achievement (Models I and II, and Submodel 5). The academic achievement is indicated by averaged scores of RCT which are considered to be more reliable than school grades. The academic achievement in turn contributes to a student's self-esteem and reduces his/her depression, anxiety, and psychological distress (Model I).

(e) Among the outcome variables, self-esteem is a major source for further changes in locus of control, mental health, and physical health (Model I and II, and Submodel 3).

(f) Initial locus of control is significantly correlated to both effective group working and constructive conflict resolution. In other words, improvements in constructive conflict resolution, as confirmed in Models I and II, and Submodels 1 and 3 combined, increases internal locus of control through reduced victimization and enhanced self-esteem; while initial internal locus of control is associated with the improvements in constructive conflict resolution. This implies the pattern predicted by Deutsch's crude law.

(g) Model IV strongly indicated that the intervention of cooperative learning at Site 2 has yielded improvements among students in effective group working. The improvements in constructive conflict resolution at Site 1 and Site 2 are suggested. But there are no statistical tests applied to these two effects. Since the findings of Model IV reflects precisely the original research design, the overall effects of the training programs are assumed.

Regarding the Interventions

The evaluative analysis in the present study is based on the modeling of factor mean structures. Model I, Model II, and the five partial models have consistently shown the induced social psychological consequences of constructive conflict resolution and possibly effective group working. Model IV indicates further that these effects may be attributable to the intervention project at Site 2 and/or Site 1. The implication is that the interventions of conflict resolution and possibly cooperative learning have generally produced real social psychological, and subsequently, educational changes at the school which are indicated by the causal estimations. In other words, Model IV bridges the findings of the theory testing to the results of the interventions, and thus provides the basis for the claims regarding the real social psychological and educational changes among the students. Therefore, the validity of Model IV is critical in our evaluation of the effectiveness of the interventions.

Alwin and Sullivan (1975) has a good summary of the problems which might be involved in an evaluational study. They classify these problems into five categories:

(1) the problem of selection processes resulting in

preintervention differences among groups;

(2) the problem of the occurrence of events unrelated to the intervention;

(3) the problem of identifying the intervening processes by which an observed intervention effect is actually produced;

(4) the problem of measurement error;

(5) the problem of differential attrition of cases due either to factors associated with selection processes, to the treatment intervention, or to postselection events unrelated to the treatment intervention.

Generally, potential challenges to the internal validity of the results of an evaluation study are those which might be legitimate factors for explaining the mean-differences on dependent measures among treatment conditions. Given the nature of our interventions and analytical methodology, Problems (2) - (4) are less of an issue compared to Problems (1) and (5). The three experimental school sites belong to one school system. Their local policies and programs are governed by the central school leadership. Under the general schoolwise administration, there could be unique events occurring at one site but not at another site. Events outside a school site but within its neighborhood may also have distinctive impact on the students at that site. However, unless an influential event is salient enough so that its occurrence has been identified and recorded, it is very difficult to control the potential confounding effect of an unknown event or intervening process. When there is no such evidence that a relevant historical event or an intervening process does exist, Problem (2) and (3) can only be treated as probabilistic; namely, potentially influencial events or intervening processes may equally likely to occur at any site. Before a better statistical approach is invented, these threats are neither justifiable nor falsifiable, quantitatively. With LISREL modeling, Problem (4) is no longer an issue.

The first and the last problems may be both viewed in the framework of selection effect. Selection is defined as the situation in which an effect may be due to the difference between the kinds of people in one experimental group as opposed to another (Cook and Campbell, 1979). No matter whether the difference in people is caused by an independent selection variable or by the interaction between the selection variable and the treatment, the consequence is exactly the same. Except that the latter case is more complicated in terms of applying control methods. When selection effect exists, the mean-differences among experimental conditions will be confounded with the effects of the pre-existing or post-intervention differences which are unrelated to the intervention.

Therefore, the main challenge to the internal validity of Model IV, the evaluative model, is the potential selection effect. Since our subjects come from three school sites which are located at different parts of the city, the population characteristics at one site might be different from those at another site. For example, our results have shown that initial locus of control has significant correlations with both cooperative learning and conflict resolution interventions. If the average level of students' internal locus of control at one site were higher than those at another site, the factor mean differences in Model IV might be confounded by the effects of locus of control.

An alternative interpretation to the intervention effects indicated in Model IV might be formed based on the differences in three-consecutive-cycle retention rate across sites. As reported previously, the students at Site 3 had much lower retention rate than those at the other two sites; while Site 2 had consistently the highest retention rate, at least for the three cycles we examined. The retention rate at Site 3 declined rapidly during those three cycles from .41 at cycle 4 to .22 at cycle 6. From our interview with the school principle after the intervention was completed, we found that the students' socio-economic background at Site 3 had been changing during the past years, and more of its students came from families with better socio-economic background. While at Site 2, students generally came from less advantaged neighborhood compared with the other two sites, particularly with Site 3. Thus it is possible that the students' socio-economic background affects their school retention rate.

Socio-economically better off students tend to be less dependent on their school. Most students in that school do not perceive themselves as likely candidates of colleges (this is an alternative high school), thus schooling is not so important to either the teachers or the students themselves. Rather, many students come to school partly because of the kind of support they may receive from the teachers or from their fellow students. For those who have a better family background, neither the schooling nor the support may be perceived as indispensable; while for those who come from a less advantaged background, the support from their teachers or fellow students, which is hard to obtain in the outside society, can be a very important reason for them to remain at the school. Thus students' socio-economic background, on one hand, affects the school retention rate at a given site; on the other hand, it fosters a special school culture in which the teachers and the students may be apt to provide the best cooperation and support to each other. If this were true, students' socio-economic background would affect both the dependent latent variable "Group" and the students' school retention, thus constitute a selection variable which might inflate the effect of cooperative learning at Site 2.

Given the quasi-experimental nature of the study, its special sampling population, and the constraints in time, resources, training and research management, numerous factors might potentially affect the accuracy of our evaluation Model. However, as we stated before, our confidence in the results of Model IV is based upon three factors: (a) the correspondence between the research design and the actual findings; (b) controlling for possible previous differences on the dependent measures by using gain scores; (c) controlling for measurement error by analyzing mean factor structures. To explicitly eliminate possible selection effects, especially potential effects of unidentified variables, more powerful statistical control methods are needed. Below, we will present a promising approach to modeling selectivity by the techniques of structural equations, which is intended to be an introduction of a follow-up study.

We may start with a simple equation to represent the selection issue, then briefly present the structural approach to modeling selectivity which may be more relevant to our study. For a regressional representation of selection effect, we may have

$$Y_{ij} = \beta_0 + \beta_1 Z_j + \beta_2 X_{ij} + e_{ij}.$$

Where Y is a criterion variable, Z stands for treatment conditions, X is a single selection variable of relevance (in an actual analytical situation, X is often expressed in its deviation form), and e is a disturbance term. The coefficient β_0 is some constant, often it is presented by μ. Y_{ij} is a score of the subject i in the treatment condition

j. This equation indicates that the criterion score of the subject i in group j is not only determined by the type of treatment he or she receives, but also affected by his or her score on the selection variable. This is a regular ANCOVA method.

Let $\beta_1 Z_j = \alpha_j$ which indicates the treatment effect in group j; furthermore, assume that the expected values of within-group disturbance terms are zero. For multi-group comparisons, we then have

$$E(Y_{ij}|j=t) = \beta_0 + \alpha_t + \beta_2{}^{(t)}E(X_{ij}|j=t), \quad t=1, 2, \ldots, g.$$

Where E(.) represents the expected value on the criterion variable of subjects in group j, t indicates the treatment condition or group. We may see that since β_0 and α_t are some constants, when β_2 and the population group means on X_{ij} are equivalent, the group differences on Y_{ij} will be solely determined by the size of α_t, namely, the effect size of treatment condition t. This reduces the equation to an ANOVA model. However, when β_2's and/or the population means are different across groups, the comparisons can only be made after the last term in the above equation have been explicitly measured and included in the equation.

As reported previously, the form of control in our analysis regarding Model IV is the gain-score method. The implication is that by using the difference scores of pre and post measures of a dependent variable, we may impose the equivalence condition on the selection variable X. Here X equals Y^* which is the pre-test score of Y. However, the gain-score adjustment can only control the effect of pre-existing difference on the same dependent variable. There are potentially other variables which might cause a selection effect but are as yet unidentified as sources of the effect.

Muthen and Joreskog (1983) have proposed an approach to modeling selectivity which may be able to account for unidentified sources of selectivity. The basic idea is similar to the above representation, however, there are two important advantages in the structural representation in this case: (a) the structural representation may treat the selection variable as a latent factor; (b) the structural equations absorb the effects of unidentified selection variables into their error terms and estimate them explicitly.

Using Muthen-Joreskog method, the selection problem may be viewed as such that there exists a threshold for a given selection variable (known or unknown), and only when a case with a value greater than the threshold on the selection variable may be observed in one particular experimental condition. For example, social disadvantage may be such a selection variable, thus a high disadvantage results in an individual being selected into an experimental condition. In multi-group comparisons, there may be different selection variables that govern the selection process for different treatment conditions. Thus, the subjects observed in one condition will be different from those in another on the selection variable(s). In an optimally randomized study, this situation may not occur. However, in a nonrandom experimental study, if the treatment groups consist of subjects having differential means on the selection variable(s), then the selection effects may be a threat to the internal validity of the experiment. Selectivity poses problems for both multi-group comparisons and consistency of the estimates of regression coefficients.

The structural equations proposed for analyzing selectivity by Muthen and Joreskog (1983) are as follows:

Controls:

$$Y^c = \mu + \beta^c X^c + \epsilon^c$$

$$T^c = s_0{}^c + s_1{}^c X^c + \delta^c$$

$$Y^c: \begin{array}{l} \text{observed if } T^c > 0 \\ \text{not observed, otherwise;} \end{array}$$

Experimentals:

$$Y^e = \mu + \alpha + \beta^e X^e + \epsilon^e$$

$$T^e = s_0{}^e + s_1{}^e X^e + \delta^e$$

$$Y^e: \begin{array}{l} \text{observed if } T^e > 0 \\ \text{not observed, otherwise.} \end{array}$$

Where Y is a single observed dependent variable (or a vector), X is an observed independent variable (or a vector), T is a unobserved selection variable (or a vector). The structural equations imply that when the selection variable T is greater than certain value (it is 0 in this case,) the case is observed. The important point with this representation is that the δ^c and δ^e include variables other than X influencing selection. Thus all the other relevant selection variables which are not specified or measured in the model may be taken into account in the error terms and estimated by this specification.

Barnow et al. (1980) and Goldberger (1979) formulated a special form of selection, where for one experimental group and one control group,

$$Y^E: \begin{cases} \text{observed if } T^E > 0 \\ \text{not observed, otherwise;} \end{cases}$$

$$Y^C: \begin{cases} \text{observed if } T^C \leq 0 \\ \text{not observed, otherwise,} \end{cases}$$

where $T^E = T^C$. They illustrated their model with the well-known Head Start Compensatory Education Program, so that Y is the post-test achievement score. Here, a single selection variable (related to family income of the child) defines group membership. The main difficulty in this type of modeling is how to implement the conditional equations in the model estimation.

From the above structural representation, it can be seen that by utilizing higher order factor analysis or path modeling, the selectivity process may be suitable to a covariance structure approach. Muthen and Joreskog (1983) has demonstrated the possibility of modeling a selection process with latent exogenous variables by using multi-group structural modeling. A set of equations are suggested for handling specification of selectivity. Unfortunately, selection processes have to be modelled specifically in each situation, no simple rules can be followed. Since the selection conditions require a model estimation to be based on a truncated distribution, LISREL program may not be readily applicable.

Our purpose of presenting a new approach to analyzing selectivity is to recommend further validation to the present findings. In many ways, the intervention research conducted in the past three years may be viewed as a landmark enterprise in the area of cooperative learning and conflict resolution. There is abundance of data collected during the study which need to be further investigated. The findings from the present study is more than encouraging, and of course, with many imperfections. Therefore, continuing efforts for further validation studies are essential in order to fully explore the significance of this large-scale intervention project.

BIBLIOGRAPHY

Abdel-Halim, A. A. (1982). Social support and managerial affective responses to job stress. Journal of Occupational Behavior, October, Vol. 3(4), 281-295.

Alwin, D. F., and Sullivan, M. J. (1975). Issues of design and analysis in evaluation research. In Bernstein, I. N., (Ed.), Validity issues in evaluation research. Beverly Hills: Sage Publications.

Amemiya, Y. (1985). On the goodness-of-fit tests for linear statistical relationships, (cited in the study of Browne and Shapiro, 1988). Technical Report, No. 10, prepared under NSF grant DMS 82-19748, Econometric Workshop, Stanford University.

Amemiya, Y., and Anderson, T. W. (1985). Asympototic chi-square tests for a large class of factor analysis models, (cited in the study of Browne and Shapiro, 1988). Technical Report, No. 13, prepared under NSF grant DMS 82-19748, Econometric Workshop, Stanford University.

Anderson, J. C. and Gerbing, D. W. (1988). Structural equation modeling in practice: a review and recommended two-step approach. Psychological Bulletin, 103, 411-423.

Anderson, T. W., Rubin, H. (1956). Statistical inference in factor analysis. Proceedings of the third Berkeley Symposium on Mathematical Statistics and Probability, (pp.111-150). Berkeley: University of California Press.

Andrews, F. M., and Robinson, J. P. (1991). Measures of subjective well-being. In Rodinson, J. P., Shaver, P. R., and Wrightsman, L. S. (Eds.), Measures of Personality and Social Psychological Attitudes. New York: Academic Press.

Andrews, G., Tennant, C., Hewson, D.. and Schonell, M. (1978). The relation of social factors to physical and psychiatric illness. American Journal of Epidemiology, 108, 27-35.

Angle, C. R., O'Brien, T. P., and McIntire, M. S. (1983). Adolescent self-poisoning: a nine-year followup. Journal of Developmental and Behavioral Pediatrics, June, Vol. 4(2), 83-87.

Arlin, M. (1975). The interaction of locus of control, classroom structure, and pupil satisfaction. Psychology in the schools, 12, 279-286.

Back, K. W. (1951). Influence through social communication. Journal of Abnormal and Social Psychology, 46, 9-23.

Baer, P. E., Garmezy, L. B., McLaughlin, R. J., and Pokorny, A. D. (1987). Stress, coping, family conflict, and adolescent alcohol use. Journal of Behavioral Medicine, October, Vol. 10(5), 449-466.

Bandura, A., Adams, N., Hardy, A. B., and Howells, G. (1980). Tests of the generality of self-efficacy theory. Cognitive Therapy Research, 39-66.

Barnow, B. S., Cain, G. G., and Goldberger, A. S. (1980). Issues in the analysis of selection bias. In Stromsdorfer, E., and Farkas, G. (Eds.), Evaluation Studies Review Annual, Vol. 5. Beverly Hills, CA: Sage.

Bentler, P. M. (1976). Multistructure statistical model applied to factor analysis. Multivariate Behavioral Research, 11, 3-25.

Bentler, P. M. (1980). Multivariate analysis with latent variables: causal modeling. Annual Review of Psychology, Vol. 31, 419-456.

Bentler, P. M. (1986). Structural modeling and Psychometrika: An historical perspective on growth and achievements. Psychometrika, Vol. 51, No. 1, 35-51.

Bentler, P. M. (1989). EQS: structural equations program manual. Los Angeles, CA: BMDP Statistical Software, Inc.

Berkman, L., and Syme, S. L. (1979). Social network, host resistence, and mortality: a nine-year follow-up study of Alameda County residents. American Journal of Epidemiology, 109(2), 186-204.

Berkowitz, L. (1957). Effects of perceived dependency relationships upon conformity to group expectations. Journal of Abnormal and Social Psychology, 55, 350-54.

Berkowitz, M. W., and Gibbs, J. C. (1985). The process of moral conflict resolution and moral development. New Direction for Child Development, September, No. 29, 71-84.

Blalock, H. M. (1964). Causal inferences in nonexperimental research. Chapel Hill: University of North Carolina.

Blalock, H. M. (1971). Causal models involving unmeasured variables in stimulus-response situations. In Blalock, H. M. (Ed), Causal models in the social sciences, (pp. 335-347). Chicago: Aldine-Atherton.

Blalock, H. M. (1979). Measurement and conceptualization problems: the major obstacle to integrating theory and research. American Sociological Review, 44, 881-894.

Blascovich, J., and Tomaka, J. (1991). Measures of self-esteem. In Rodinson, J. P., Shaver, P. R., and Wrightsman, L. S. (Eds.), Measures of Personality and Social Psychological Attitudes. New York: Academic Press.

Blazer, D. G. (1982). Social support and mortality in an elderly community population. American Journal of Epidemiology, 115, 684-694.

Block, J., and Thomas, H. (1955). Is satisfaction with self a measure of adjustment? Journal of Abnormal and Social Psychology, 51, 254-59.

Bloom, B., Asher, S., White, S. (1978). Marital disruption as a stressor: a review and analysis. Psychological Bulletin, 85, 867-894.

Blumberg, P., and Flaherty, J. A. (1985). The influence of noncognitive variables in student performance. Journal of Medical Education, September, Vol. 60(9), 721-23.

Bollen, K. A. (1989). Structural equations with latent variables. New York: Wiley.

Boneau, C. A. (1959). The effects of violations of assumptions underlying the t test. Psychological Bulletin, 57, 49-64.

Bonoma, T., Tedeschi, J., and Helm, B. (1974). Some effects of target cooperation and reciprocated promises on conflict resolution. Sociometry, 37, 251-261.

Bowlby, J. (1969). Attachment and loss: Volume 1. London: Hogarth Press.

Bowlby, J. (1973). Attachment and loss: Volume 2. London: Hogarth Press.

Bowlby, J. (1980). Attachment and loss: Volume 3. New York: Basic Books.

Bradburn, N. M. (1968). The structure of psychological well-being. Aldine Publishing Co., Chicago.

Brickman, P., Rabinowitz, V. C., Karuza, J., Coates, D., Cohn, E., and Kidder, L. (1982). Models of helping and coping. American Psychologist, 37, 368-384.

Bridgeman, D. (1977). The influence of cooperative, interdependent learning on role taking and moral reasoning: a theoretical and empirical field study with fifth grade students. Unpublished Doctoral dissertation, University of California, Santa Crutz.

Brody, D. S. (1980). Psychological distress and hypertension control. Journal of Human Stress, 6, 2-6.

Brook, R. H., et al. (1980). Conceptualization and measurement of physiologic health for adults, (Cited in Ware et al.'s). Santa Monica, CA: Rand Corporation.

Brown, G. W., and Harris, T. (1978). Social origins of depression: a study of psychiatric disorders in women. New York, Free Press.

Browne, M. W., and Shapiro, A. (1988). Robustness of normal theory methods in the analysis of linear latent variate models. British Journal of Mathematical and Statistical Psychology, November, Vol. 41(2), 193-208.

Browne, M. W. (1985). Robustness of normal theory tests of fit of factor analysis and related models against non-normally distributed common factors, (cited in the study of Browne and Shapiro, 1988), prepared at the Fourth European Meeting of the Psychometric Society and the Classification Societies, Cambridge, 2-5 July.

Browne, M. W. (1987). Robustness of statistical inference in factor analysis and related models. Biometrika, 74, 375-384.

Bucke, R., and Parke, R. D. (1972). Behavioral and physiological response to the presence of a friendly or neutral person in two types of stressful situations. Journal of Personality and Social Psychology, 24, 143-153.

Burt, R. S. (1976). Interpretational confounding of unobserved variables in structural equation models. Sociological Methods and Research, August, 3-52

Caldwell, R. A., Pearson, J. L., and Chin, R. J. (1987). Stress moderating effects: social support in the context of gender and locus of control. Personality and Social Psychology Bulletin, March, Vol. 13(1), 5-17.

Campbell, D. T., and Stanley, J. C. (1966). Experimental and quasi-experimental designs for research. Chicago: Rand McNally.

Caplan, R. D., Robinson, E. A. R., French, J. R. P., Caldwell, J. R., and Shinn, M. (1976). Adhering to medical regimens: pilot experiments in patient education and social support. Ann Arbor, Michigan: Institute for Social Research.

Caplan, R. D., Harrison, R. V., Wellons, R. V., and French, J. R. P. (1980). Social support and patient adherence. Ann Arbor, Michigan: Institute for Social Research.

Cassel, J. (1976). The contribution of the social environment to host resistence. American Journal of Epidemiology, 104, 107-123.

Cauce, A. M. (1986). Social networks and social competence: exploring the effects of early adolescent friendships. American Journal of Community Psychology, December, Vol. 14(6), 607-628.

Chen, H., and Rossi, P. H. (1983). Evaluating with sense: the theory-driven approach. Evaluation Review, Vol. 7, No. 3, June, 283-302.

Chen, H., and Rossi, P. H. (1987). The theory-driven approach to validity. Evaluation and Program Planning, Vol. 10, 95-103.

Cobb, S. (1976). Social support as a moderator of life stress. Psychosomatic Medicine, 38, 300-314.

Cobb, S., and Kasl, S. V. (1977). Termination: the consequence of job loss. Publication NIOSH 77-224. Washington, DC: Department of Health, Education and Welfare.

Cohen, S., and Wills, T. H. (1985). Stress, social support and buffering hypothesis. Psychological Bulletin, 98, 310-357.

Cohen, S., and Syme, L. (1985). Social support and Health. New York: Academic Press.

Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D., and York, R. L. (1966). Equality of education Opportunity. Washington, DC: United States Government Printing Office.

Collins, B. A., Whalen, C. K., and Henker, B. (1980). Ecological and pharmacological influences of behaviors in the classroom: the hyperkinetic syndrome. In

Salziner, S., Antrobus, J., and Glick, J. (Eds.), The eco-system of the "sick child." New York: Academic Press.

Comstock, L., and Slome, C. (1973). A health survey of students 1: prevalence of Problems. Journal of American College Health Association, 22, 150-55.

Conn, M. K., and Christopher, P. (1989). Social support: seek and ye shall find. Journal of Social and Personal Relationships, August, Vol. 6(3), 345-358.

Cook, T. D., and Campbell, D. T, (1979). Quasi-experimentation: design and analysis issues for field settings. Boston: Houghton Mifflin Company.

Coopersmith, S. (1967). The antecedents of self-esteem. San Francisco: W. H. Freeman & Co.

Corcoran, K. J., Robbins, S. P., and Hepler, S. E. (1987). Support systems and scholarly productivity, Journal of Applied Social Sciences, Spring-Summer, Vol. 11(2), 230-242.

Cotton, J. L., and Cook, M. S. (1982). Meta-analysis and the effects of various reward systems: some different conclusions from Johnson et al.. Psychological Bulletin, Vol. 92, No. 1, 176-83.

Cramer, D. (1988). Self-esteem and facilitative close relationships: a cross-lagged panel correlation analysis. British Journal of Social Psychology, June, Vol. 27(2), 115-126.

Crandall, R. (1976). The measurement of self-esteem and related constructs. In Robinson, J. P., and Shaver, P. R. (1976), Measures of Social Psychological Attitudes. The University of Michigan, Ann Arbor, Michigan: Institute for Social Research.

Crombag, H. (1966). Cooperation and competition in means-interdependent triads: a replication. Journal of Personality and Social Psychology. 4(6), 692-95.

Crouse, R. H. (1985). Using peer network theory with a residential program for Chicano students. Journal of College Student Personnel, November, Vol. 26(6), 549-550.

Cummins, R. (1988). Perception of social support, receipt of supportive behaviors, and locus of control as moderators of the effects of chronic stress. American Journal of Community Psychology, October, Vol. 16(5), 685-700.

Darley, J. M., and Goethals, G. R. (1980). People's analysis of the causes of ability-linked performances. In Berkowitz, L. (Ed.), Advances in experimental social psychology, Vol. 13. New York: Academic Press.

Dean, A., and Lin, N. (1977). The stress-buffering role of social support: problems and prospects for systematic investigation. Journal of Nervous and Mental Disease, 165, 403-417.

Deaux, K. (1976). Sex: a perspective on the attribution process. In Harvey, J. H., Ickes, W. J., and Kidd, R. F. (Eds.), New directions in attribution research. Hillsdale, NJ: Erlbaum.

Delongis, A., Folkman, S., and Lazarus, R. S. (1988). The impact of daily stress on health and mood: psychological and social resources as mediators. Journal of Personality and Social Psychology, March, Vol. 54(3), 486-495.

Depner, C. E., and Wethington, E. (1984). Social support: mathodological issues in design and measurement. Journal of Social Issues, Vol. 40, No. 4, 37-54.

Deutsch, M. (1948). The effects of cooperation and competition upon group process. Dissertation.

Deutsch, M. (1949a). An experimental study of the effects of co-operation and competition upon group process. Human Relations, Vol. II, No. 2, April, 129-152..

Deutsch, M. (1949b). An experimental study of the effects of co-operation and competition upon group process. Human

Relations, Vol. II, No. 3, July, 199-231.

Deutsch, M. (1958). Trust and suspicion. Journal of Conflict Resolution, 2, 265-279.

Deutsch, M. (1960). The effect of motivational orientation upon trust and suspicion. Human Relations, 13, 123-139.

Deutsch, M. (1962). Cooperation and trust: some theoretical notes. In Jones, M., (Ed.), Nebraska symposium on motivation. Lincoln: University of Nabraska Press.

Deutsch, M. (1968). The effects of cooperation and competition upon group process. In D. Cartwright & A. Zander (Eds.), Group Dynamics (Pp. 461-482). New York: Harper & Row, Publishers.

Deutsch, M. (1973). The resolution of conflict: constructive and destructive processes. New Haven: Yale University Press.

Deutsch, M. (1985). Distributive justice: a social-psychological perspective. New Haven: Yale University Press.

Deutsch, M. (1987). A framework for teaching conflict resolution in the schools. Paper presented at the Negotiations in Organizations Conference in Mt. Sterling, Ohio, April.

Deutsch, M., Boardman, S. K., Doucette-Gates, A., & Raider, E. (1988). Research on the effects of training in conflict resolution and cooperation. Interim Progress Report: Project II.8.

Duncan, O. D. (1966). Path analysis: Sociological examples. American Journal of Sociology, 72, 1-16.

Duncan, O. D., Haller, A. O., and Portes, A. (1968). Peer influence on aspirations: a reinterpretation. American Journal of Sociology, 74, 119-137.

Durbin, J. (1954). Errors in variables. Review of the

International Statistics Institute, 1, 23-32.

Dweck, C. S. (1975). The role expectations and attributions in the alleviation of learned helplessness. Journal of Personality and Social Psychology, 31, 674-685.

Dweck, C. S., and Bush, E. S. (1976). Sex differences in learned helplessness: I. differential debilitation with peer and adult evaluators. Developmental Psychology, 12, 147-156.

Dweck, C. S., Davidson, W., Nelson, S., and Enna, B. (1978). Sex differences in learned helplessness: II. The contingencies of evaluative feedback in the classroom and III. An experimental analysis. Developmental Psychology, 14, 268-276.

Eckenrode, J. (1983). The mobilization of social supports: some individual constraints. American Journal of Community Psychology, October, Vol. 11(5), 509-528.

Egbert, L. D., Battit, G. E., Welch, C. E., and Bartlett, M. K. (1964). Reduction of post-operative pain for encouragement and instruction of patients. New England Journal of Medicine, 270, 825-27.

Epley, S. W. (1974). Reduction of the behavioral effects of adversive stimulation by the presence of companions. Psychological Bulletin, 81, 271-283.

Fay, A. (1970). The effects of cooperation and competition on learning and recall, (Cited in the study of Johnson, D. W., and Johnson, R. (1989)). Unpublished master's thesis, George Peabody College.

Fazio, A. F. (1977). A concurrent validational study of the NCHS General Well-Being Schedule. (Dept. of H. E. W. Publ. No. HRA-78-1347). Hyattsville, MD: National Center for Health Statistics.

Feather, N. T., and Simon, J. G. (1973). Fear of success and causal attributions for outcome. Journal of Personality, 41, 515-542.

Fisher, J. D., and Nadler, A. (1974). Effect on similarity between donor and recipient on recipients reactions to aid. Journal of Applied Social Psychology, 4, 230-243.

Fitts, W. (1972). The self concept and psychopathology, (Monograph 4). Nashville, Tenn.: Counselor Recordings and Tests.

Fornell, C. (1982). A second generation of multivariate analysis -- An overview. In Fornell, C. (Ed.), A second generation of multivariate analysis, (Vol. 1), (pp. 1-21). New York: Praeger Publishers.

Fornell, C., and Larcker, D. F. (1982). Structural equation models with unobservable variables and measurement error: algebra and statistics. In Fornell, C. (Ed.), A second generation of multivariate analysis, (Vol. 2), (pp. 331-347). New York: Praeger Publishers.

Franklin, R. D. (1963). Youth's experiencies about internal versus external control of reinforcement related to N variables, (cited in Lefcourt's study, 1982). Dissertation Abstracts, 24, 1684.

French, J. (1951). Group productivity. In Guetzkow, H. (Ed.), Groups, leadership and men, (pp. 44-55). Pittsburgh: Carnegie Press.

Fusilier, M. R., Ganster, D. C., and Mayes, B. T. (1987). Effects of social support, role stress, and locus of control on health. Journal of Management, Fall, Vol. 13(3), 517-528.

Ganellen, R. J., and Blaney, P. H. (1984). Hardiness and social support as moderators of the effects of life stress. Journal od Personality and Social Psychology, July, Vol. 47(1), 156-163.

Gavin, J. F., Axelrod, W. L. (1977). Managerial stress and strain in a mining organization. Journal of Vocational Behavior, August, Vol. 11(1), 66-74.

Geer, J. H., Davison, G. C., and Gatchel, R. I. (1970). Reduction of stress in humans through non-vertical perceived control of aversive stimulation. Journal of

Personality and Social Psychology, 16, 731-38.

Glass, D. C. (1977). Stress, behavior patterns, and coronary disease. American Scientist, 65, 177-87.

Glass, D. C., Singer, J. E., Leonard, H. S., Krantz, D., Cohen, S., and Cummings, H. (1973). Perceived control of aversive stimulation and the reduction of stress responses. Journal of Personality, 41, 577-595.

Goldberger, A. S. (1964). Econometric theory. New York: Wiley.

Goldberger, A. S. (1970). On Boudon's method of linear causal analysis. American Sociological Review, 35, 97-101.

Goldberger, A. S. (1979). Methods for eliminating selection bias. Department of Economics, University of Wisconsin, Madison.

Gore, S. (1978). The effect of social support in moderating the health consequences of unemployment. Journal of Health and Social Behavior, 19, 157-165.

Green, S. G., and Mitchell, T. R. (1979). Attributional processes of leaders in leader-member interactions. Organizational Behavior and Human Performance, 23, 429-458.

Greenley, J., and Mechanic, D. (1976). Social selection in seeking help for psychological problems. Journal of Health and Social Behavior, 17, 249-262.

Grossack, M. (1954). Some effects of cooperation and competition on small group behavior. Journal of Abnormal and Social Psychology, 49, 341-48.

Guttman, L. (1952). Multiple group methods for common factor analysis: Their basis, computation, and interpretation. Psychometrika, 17, 209-222.

Harpin, P. M. (1980). Individual variability in behavior within a person-environment interactional design, (cited in Lefcourt's study, 1982). Unpublished dissertation, Arizona State University.

Harrison, F. I. (1968). Relationship between home background, school success, and adolescent attitudes. Merrill-Palmer Quarterly of Behavior Development, 14, 331-344.

Hartley, D., Roback, H. B., and Abramovitz, S. I. (1976). Deterioration effects in encounter groups. American Psychologist, 31, 247-255.

Havlicek, L. L., and Peterson, N. L. (1974). Robustness of the t test: A guide for researchers on effect of violations of assumptions. Psychological Reports, 34, 1095-1114.

Haynes, N. M., and Comer, J. (1990). A comparison of learning and motivation among high school students: correction. Psychology in Schools, July, Vol. 27(3), 275.

Haynes, N. M. (1990). A comparison of learning and motivation among high school students: correction. Psychology in Schools, April, Vol. 27(2), 163-171.

Hempel, C. G. (1965). Studies in the logic of confirmation. In Hempel, C., Aspects of scientific explanation and other essays. New York: Free Press.

Hempel, C. G. (1966). Philosophy of natural science. Englewood Cliffs, NJ: Prentice-Hall.

Henderson, S., Byrne, D., Duncan-Jones, P., Scott, R., and Adcock, S. (1980). Social relationships, adversity and neurosis: a study of associations in a general population sample. British Journal of Psychiatry, 136, 574-583.

Hibbard, J. H. (1985). Social ties and health status: an examination of moderating factors. Health Education Quarterly, Spring, Vol. 12(1), 23-34.

Hoffman, M. A., Ushpiz, V., and Levy-Shiff, R. (1988). Social support and self-esteem in adolescence. Journal of Youth and Adolescence, August, Vol. 17(4), 307-316.

Hoge, D., Smit, E. K., and Hanson, S. L. (1990). School experiences predicting changes in self-esteem of sixth- and seventh-grade students. Journal of Educational Psychology, Vol. 82, No. 1, 117-127.

Howard, S., and Kubis, J. (1964). Ego identity and some aspects of personal adjustment. Journal of Psychology, 58, 459-466.

Howell, D. C. (1982). Statistical Methods for Psychology. Boston: Duxbury Press.

Huang, H. C., Hwang, K. K., and Ko, Y. H. (1983). Life stress, attribution style, social support and depression among university students. Acta Psychologica Taiwanica, June, Vol. 25(1), 31-47.

Husaini, B. A., and von Frank, A. (1985). Life events, coping resources, and depression: a longitudinal study of direct, buffering, and reciprocal effects. Research in Community and Mental Health, Vol. 5, 111-136.

Jackson, S. E. (1983). Participation in decision making as a strategy for reducing job-related strain. Journal of Applied Psychology, February, Vol. 68(1), 3-19.

James, W. H. (1965). The application of social learning theory to educational processes, (cited in Lefcourt's study, 1982). Paper presented at a meeting of the Society for Research in Child Development, Minneapolis.

James, L. R., and Brett, J. M. (1984). Mediators, moderators, and tests for mediation. Journal of Applied Psychology, Vol. 69, No. 2, 307-321.

Jason, L. A., Tabon, D., Tait, E., Iacono, G., Goodman, D., Watkins, P., and Huggins, G. (1988). The emergence of the inner-city self-help center. Journal of Community Psychology, July, Vol. 16(3), 287-295.

Jecker, J., and Landy, D. (1969). Liking a person as a function of doing him a favor. Journal of Human Relations, 17(4), 630-40.

Jenkins, C. D. (1979). Psychosocial modifiers of response to stress. Journal of Human Stress, 5, 3-15.

Johnson, D. W., and Ahlgren, A. (1976). Relationship between students' attitudes about cooperative learning and competition and attitudes toward schooling. Journal of Educational Psychology, 68, 92-102.

Johnson, D. W., and Johnson, S. (1972). The effects of attitude similarity, expectations of goal facilitation on interpersonal attraction. Journal of Experimental and Social Psychology, 8, 197-206.

Johnson, D. W., and Johnson, R. T. (1983). Social interdependence and perceived academic and personal support in the classroom. Journal of Social Psychology, 120, 77-82.

Johnson, D. W., and Johnson, R. (1984). Cross-ethnic relationships: Impact of intergroup cooperation and intergroup competition. Journal of Educational Research, 78(2), 75-79.

Johnson, D. W., and Johnson, R. (1987). Learning together and alone: Cooperation, competition, and individualization, (2nd. ed.). Englewood Cliffs, NJ: Prentice-Hall.

Johnson, D. W., and Johnson, R. T. (1989). Cooperation and competition: theory and research. Minnesota: Interaction Book Company.

Johnson, D. W., Johnson, R. T., and Anderson, D. (1978). Student cooperative, competitive, and individualistic attitudes and attitudes toward schooling. Journal of Psychology, 100, 183-199.

Johnson, D. W., Johnson, R. T., and Anderson, D. (1983). Social interdependence and classroom climate. Journal of Psychology, 114, 135-142.

Johnson, D. W., Johnson, R. T., Buckman, L., and Richards, P. (1986). The effect of prolonged implementation of cooperative learning on social support within the classroom. Journal of Psychology, 119, 405-411.

Johnson, D. W., Johnson, R. T., and Scott, L. (1978). The effects of cooperative and individualized instruction on student attitudes and achievement. Journal of Social Psychology, 104, 207-216.

Johnson, D. W., Johnson, R., Tiffany, M., and Zaidman, B. (1983). Are low achievers disliked in a cooperative situation? A test of rival theories in a mixed-ethnic situation. Contemporary Educational Psychology, 8, 189-200.

Johnson, D. W., Maruyama, G., Johnson, R., Nelson, D., and Skon, L. (1981). Effects of cooperative, competitive and individualistic goal structures on achievement: A meta-analysis. Psychological Bulletin, 89, 47-62.

Johnson, D. W., Maruyama, G., and Johnson, R. T. (1982). Separating ideology from currently available data: a reply to Cotton and Cook and McGlynn. Psychological Bulletin, Vol. 92, No. 1, 186-92.

Jones, S. (1973). Self and interpersonal evaluations: esteem versus consistency theories. Psychological Bulletin, 79, 185-199.

Joreskog, K. G. (1967). Some contributions to maximum likelihood factor analysis. Psychometrika, 32, 443-482.

Joreskog, K. G. (1969). A general approach to confirmatory maximum likelihood factor analysis. Psychometrika, 34, 183-202.

Joreskog, K. G. (1973). A general method for estimating a linear structural equation system. In Goldberger, A. S., and Duncan, O. D. (Eds.), Structural equation models in the social sciences, (85-112). New York: Academic Press.

Joreskog, K. G. (1982). Reply. In Fornell, C. (Ed.), A second generation of multivariate analysis, (Vol. 1), (pp. 1-21). New York: Praeger Publishers.

Joreskog, K. G., and Sorbom, D. (1979). Advances in factor analysis and structural equation models. Cambridge, MA.: Abt Books.

Joreskog, K. G., and Sorbom, D. (1989). LISREL 7: A guide to the program and application, (2nd. ed.). SPSS INC.

Kahn, R. L., and Antonucci, T. C. (1981). Convoys of social support: a life course approach. In Kiesler, S., Morgan, J., and Oppenheimer, V. (Eds.), Aging: social change. New York: Academic Press.

Kahn, R. (1981). Work and Health. New York: Wiley.

Kaplan, B. H., Cassel, J. C., and Gore, S. (1977). Social Support and Health. Medical Care, 25, 47-58.

Kaplan, B. H., Cassel, J. C., and Gore, S. (1979). Social support and health. In Jaco, E. G. (Ed.), Patients, physicians, and illness: a sourcebook in behavioral science and health. New York: Free Press.

Kashani, J. H., and Sheppard, J. A. (1990). Aggression in adolescents: the role of social support and personality. Canadian Journal of Psychiatry, May, Vol. 35(4), 311-315.

Khattri, N. (1991). An assessment of the social validity of cooperative learning and conflict resolution programs in an alternative inner city high school. Unpublished Project Report, International Center for Cooperation and Conflict Resolution, Teachers College, Columbia University.

Katz, I. (1967). The socialization of academic motivation in minority group children. In Levine, D. (Ed.), Nebraska Symposium on motivation. Lincoln, NE: University of Nebraska Press.

Keesling, J. W. (1972). Maximum likelihood approaches to causal analysis. Ph.D. dissertation. Department of Education, University of Chicago.

Kelley, H. H. (1979). Personal relationships: the structure and processes. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Kessler, R. C. (1982). Life events, social support, and mental health. In Walter, R. G. (Ed.), Deviance and mental illness. Beverly Hills, CA: Sage.

Kessler, R. C. (1983). Methodological issues in the study of psychological stress. In Kaplan, H. B. (Ed.), Psychological stress: trends in research and theory, (pp. 267-341). New York: Academic Press.

Kessler, R. C., and McLeod, J. D. (1985). Social support and mental health in community samples. In Cohen, S., and Syme, S. L. (Eds.), Social support and health, (pp. 219-240). New York: Academic Press.

Kuhn, T. S. (1970). The structure of scientific revolutions, (2nd. ed., enlarged). Chicago: The University of Chicago Press.

Langer, E. J., and Rodin, J. (1976). The effects of choice and enhanced personal responsibility for the aged: a field experiment in an institutional setting. Journal of Personality and Social Psychology, 34, 191-98.

La Rocco, J. M., House, J. S., and French, J. R. P. (1980). Social support, occupational stress, and health. Journal of Health and Social Behavior, 21, 202-218.

La Rocco, J. M., and Jones, A. P. (1978). Co-worker and leader support as moderators of stress-strain relationships in work situations. Journal of Applied Psychology, October, Vol. 63(5), 629-634.

Lawley, D. N. (1967). Some new results in maximum likelihood factor analysis. Proceedings of the Royal Society of Edinburgh, A67, 256-264.

Leavy, R. (1983). Social support and psychological disorder: a review. Journal of Community Psychology.

Lefcourt, H. M. (1981). Locus of control: current trends in theory and research. Hillsdale, NJ: Erlbaum.

Lefcourt, H. M. (Ed.) (1984). Research with the locus of control construct (Vol. 3). New York: Academic Press.

Lefcourt, H. M., Martin, R. A., and Saleh, W. E. (1984). Locus of control and social support: interactive moderators of stress. Journal of Personality and Social Psychology, August, Vol. 47(2), 378-389.

Lerner, M., and Simmons, C. (1966). Observer's reaction to the "innocent victim": compassion or rejection? Journal of Personality and Social Psychology, 4, 203-210.

Lessing, E. E. (1969). Racial differences in indices of ego functioning relevant to academic achievement. Journal of Genetic Psychology, 115, 153-167.

Levenson, H. (1981). Differentiating among internality, powerful others, and chance. In Lefcourt, H. M. (Ed.), Research with the locus of control construct (Vol. 1, pp. 15-63). New York, Academic Press.

Leventhal, H. (1980). Applied social psychological research: the salvation of substantive social psychological theory. In Kidd, R., and Saks, M. (Eds.), Advances in applied social psychology. Hillsdale, N.J.: Erlbaum.

Lieberman, M. A., and Mullen, J. T. (1978). Does help help? The adaptive consequences of obtaining help from professionals and social networks. American Journal of Community Psychology, 6, 499-517.

Lin, N., Simone, R. L., Ensel, W. M., and Kuo, W. (1979). Social support, strssful life events and illness: a model and an empirical test. Journal of Health and Social Behavior, 20: 108-119.

Lipsey, M. W., and Pollard, J. A. (1989). Driving toward theory in program evaluation: more models to choose from. Evaluation and Progran Planning, Vol. 12, 317-328.

Little, V. L., and Kendall, P. C. (1978). Note on locus of control and academic achievement in institutionalized delinquents. Journal of Abnormal Child Psychology, 6, 281-283.

Lord, F. M., and Novick, M. R. (1968). Statistical theories of mental test scores. Reading, MA: Addison-Wesley.

Lynch, J. L. (1977). The broken heart: the medical consequences of loneliness. New York: Basic Books.

Madansky, A. (1959). The fitting of straight lines when both variables are subject to error. Journal of the American Statistical Association, 54, 173-205.

Mallinckrodt, B. (1989). Social support and effectiveness of group therapy. Journal of Counseling Psychology, April, Vol. 36(2), 170-175.

Mallinckrodt, B. (1988). Student retention, social support, and dropout intention: comparison of black and white students. Journal of College Student Development, January, Vol. 29(1), 60-64.

Marcelissen, F. H., Winnubst, J. A., Buunk, B., and de Wolff, C. J. (1988). Social support and occupational stress: a causal analysis. Social Science and Medicine, Vol. 26(3), 365-373.

Marsh, H. W., and Parker, J. W. (1984). Determinants of student self-concept: is it better to be a relatively large fish in a small pond even if you don't learn to swim as well? Journal of Personality and Social Psychology, 47, 213-231.

Marwell, G., and Hage, J. (1970). The organization of role relationships: a systematic description. Americal Sociological Review, 35, 884-900.

Maruyama, G., Rubin, R. A., and Kingsbury, G. G. (1981). Self-esteem and educational achievement : independent constructs with a common cause? Journal of Personality and Social Psychology, 40, 962-975.

McDonald, R. P. (1985). Factor analysis and related methods.

Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

McGlynn, R. P. (1982). A comment on the meta-analysis of goal structures. Psychological Bulletin, Vol. 92, No. 1, 184-85.

Mitchell, R., Billings, A., and Moos, R. (1982). Social support and well-being: implications for prevention programs. Journal of Primary Prevention, 3, 77-98.

Mitchell, R. E., Mattews, J. R., Grandy, T. G., and Lupo, J. V. (1983). The question of stress among first-year medical students. Journal of Medical Education, May, Vol. 58(5), 567-372.

Mitchell, V. (1991). A qualitative study of training on conflict resolution and cooperative learning in an alternative high school. Project Report, International Center of Cooperation and Conflict Resolution, Teachers College, Columbia University.

Moos, R., and Van Dort, B. (1976). Student physical symptoms and the social climate of college living groups, (cited in Rodin's study, 1985). Unpublished manuscript. Social Ecology Laboratory, Department of Psychiatry and Behavioral Sciences, Stanford University.

Mulaik, S. (1987). A brief history of the philosophical foundations of exploratory factor analysis. Multivariate Behavioral Research, 22, 267-305.

Mullen, B., and Suls, J. (1982). Know thyself: stressful life changes and the ameliorative effect of private self-consciousness. Journal of Experimental Social Psychology, 18, 43-55.

Mutten, B., and Joreskog, K. G. (1983). Selectivity problems in quasi-experimental studies. Evaluation Review, Vol. 7, No. 2, April, 139-174.

Newton, I. (1953). Newton's philosophy of nature: selections from his writings. New York: Haffner.

Newcomb, M. D. (1990). Social support and personal charateristics: a developmental and interactional perspective. Journal of Social and Clinical Psychology, Spring, Vol. 9(1), 54-68.

Nisbett, R. E., and Ross, L. (1980). Human inference: strategies and shortcomings of social judgment. Englewood Cliffs, N.J.: Prentice-Hall.

Norbeck, J. S., and Tilden, V. P. (1983). Life stress, social support and emotional diseqilibrium in complications of pregnancy: a prospective, multivariant study. Journal of Health and Social Behavior, 24(1), 30-45.

Norton, D. W. (1953).The approximate randomization test as an alternative to the F test in analysis of variance, (cited in Still, A. W., and White, A. P., 1981). British Journal of Mathematics and Statistical Psychology, 34, 243-52.

Nowicki, S., and Roundtree, J. (1971). Correlates of locus of control in secondary school age students, (cited in Lefcourt's study, 1982). Unpublished manuscript, Emory University.

Nuckolls, K. B., Cassel, J., and Kaplan, B. H. (1972). Psychosocial assets, life crisis and the prognosis of pregnancy. American Journal of Epidemiology, 95, 431-441.

Nunnally, J. C. (1978). Psychometric theory, (2nd. ed.). New York: McGraw-Hill.

Oei, T. I., and Zwart, F. M. (1990). The role and development of some methodological questions in life event, social support, and depression research. Stress Medicine, April-June, Vol. 6(2), 127-132.

O'Malley, P., and Bachman, J. (1979). Self-esteem and education: sex and cohort comparisons among high school seniors. Journal of Personality and Social Psychology, 37, 1153-1159

Pareek, U., and Narendra, D. (1977). Effect of partner's response and communication on competitive and cooperative game behavior. Psychologia: An

International Journal of Psychology in the Orient, March, Vol. 20(1) 38-48.

Pearlin, L., and Schooler, C. (1978). The structure of coping. Journal of Health and Social Behavior, 19, 2-21.

Pfannenstiel, J. C., Seltzer, D. A. (1989). New parents as teachers: evaluation of an early parent eduaction program. Early Childhood Research Quarterly, March, Vol. 4(1), 1-18.

Purkey, W. (1970). Self-concept and school achievement. Englewood Cliffs, NJ: Prentice-Hall.

Pyskoty, C. E., Richman, J. A., and Flaherty, J. A. (1990). Psychosocial assets and mental health of minority medical students. Academic Medicine, September, Vol. 65(9), 581-585.

Quick, J. D., Horn, R. S., and Quick, J. C. (1986). Health consequences of stress. Journal of Organizational Behavior Management, Fall-Winter, Vol. 8(2), 19-36.

Raven, B. H., and Eachus, H. T. (1963). Cooperation and competition in means-interdependent triads. Journal of Abnormal and Social Psychology, 67, 307-16.

Reynolds, W. M. (1988). Measurement of academic self-concept in college students. Journal of Personality Assessment, 52, 223-240.

Rindskopf, D. (1983). Parameterizing inequality constraints on unique variances in linear structural models. Psychometrika, 48, 73-83.

Rindskopf, D. (1984). Structural equation models: empirical identification, Heywood cases, and related problems. Sociological Methods and Research, 13, 109-119.

Rivicki, D. A., and May, H. J. (1985). Occupational stress, social support, and depression. Health Psychology, Vol. 4(1), 61-77.

Rodin, J. (1977). Research on eating behavior and obesity: where does it fit in personality and social psychology? Personality and Social Psychology Bulletin, 3, 333-355.

Rodin, J. (1985). The application of social psychology. In Lindzey, G., and Aronson, E. (Eds.), The handbook of social psychology, (3rd. ed.), New York: Random House.

Rodin, J., and Janis, I. L. (1979). The social power of health-care practitioners as agents of change. Journal of Social Issues, 35, 60-81.

Rodin, J., and Langer, E. J. (1977). Long term effects of a control-relevant intervention with the institutionalized aged. Journal of Personality and Social Psychology, 35, 897-902.

Rosenberg, M. (1965). Society and the adolescent self-image. Princeton, NJ: Princeton University Press.

Rosenberg, M. (1979). Conceiving the self. New York: basic Books.

Rothenberg, T. J. (1971). Identification in parametric models. Econometrica, Vol. 39, No. 3, 577-591.

Roy, C. M., and Fugua, D. R. (1983). Social support systems and academic performance of single-parent students. School Counselor, January, Vol. 30(3), 183-192.

Ruby, T., and Law, R. (1982). School dropouts--they are not what they seem to be. Children and Youth Services Review, Vol. 4(3), 279-291.

Sandler, I. (1980). Social support resources, stress, and maladjustment of poor children. American Journal of Community Psychology, 8, 41-52.

Sanford, N. (1970). Whatever happened to action research? Journal of Social Issues, 26, 3-23.

Sarason, I., Levine, H., Basham, R., and Sarason, B. (1983). Assessing social support: The social support questionnare. Journal of Personality and Social Psychology, 44, 127-139.

Schaubroeck, J. Cotton, J. L., and Jennings, K. R. (1989). Antecedents and consequences of role stress: a covariance structure analysis. Journal of Organizational Behavior, January, Vol. 10(1), 35-58.

Scheirer, M. A., and Kraut, R. E. (1979). Increasing educational achievement via self concept change. Review of Educational Research, 49, 131-150.

Schmale, A., and Iker, H. (1966). The psychological setting of uterine cervical cancer. Annals of New York Academy of Sciences, 125, 807-813.

Schulz, R. (1976). The effects of control and predictability on the psychological and physical well-being of the institutionalized aged. Journal of Personality and Social Psychology, 33, 563-73.

Seers, A., McGee, G., Serey, T., and Graen, G. (1983). The interaction of job stress and social support: a strong influence investigation. Academy of Management Journal 26, 273-284.

Seta, J., Paulus, P., and Schkade, J. (1976). Effects of group size and proximity under cooperative and competitive conditions. Journal of Personality and Social Psychology, 34, 47-53.

Sharan, S. (1980). Cooperative learning in teams: Recent methods and effects on achievement, attitudes, and ethnic relations. Review of Educational Research, 50, 241-272.

Silber, E., and Tippett, J. (1965). Self-esteem: clinical assessment and measurement validation. Psychological Reports, 16, 1017-1071.

Silver, R. L., and Wortman, C. B. (1980). Coping with undesirable life events. In Garber, J., and Seligman, M. E. P. (Eds.), Human helplessness. New York: Academic Press.

Simon, H. A. (1954). Spurious correlaiton: A causal interpretation. Journal of the American Statistical Association, 49, 467-479.

Skaalvik, E. M., and Hagtvet, K. A. (1990). Academic achievement and self-concept: an analysis of causal predominance in a developmental perspective. Journal of Personality and Social Psychology, Feb. Vol. 58(2), 292-307.

Slavin, R. (1983). Cooperative learning. New York, Longman.

Slavin, R. (1984). Meta-analysis in education: How has it been used? Educational Researcher, 13(8), 6-15.

Slavin, R., and Karweit, N. (1981). Cognitive and affective outcomes of an intensive student team learning experience. Journal of Experimental Education, 50, 29-35.

Smith, M. L., and Glass, G. V. (1977). Meta-analysis of psychotherapy outcome studies. American Psychologist, 32, 752-760.

Sober, E. (1988). Reconstructing the past. Cambridge, MA: A Bradford Book, the MIT Press.

Spearman, C. (1904). "General intelligence" objectively determined and measured . American Journal of Psychology, 15, 201-292.

Steil, J. M. (1980). Efficacy and the response to injustice by relatively advantaged and disadvantaged persons. Ph.D. Dissertation, Teachers College, Columbia University.

Steil, J. M. (1983). The response to injustice: effects of varying levels of social support and position of advantage or disadvantage. Journal of Experimental Social Psychology, 19, 239-253.

Steil, J. M., Tuchman, B., and Deutsch, M. (1978). An exploratory study of the meanings of injustice and frustration. Personality and Social Psychology

Bulletin, 4, 393-398.

Stephens, M. W. (1973). Parent bahavior antecedents, cognitive correlates and multidimensionality of locus of control in young children, (cited in Lefcourt's study, 1982). Paper presented at American Psychological Association Convention, Montreal.

Still, A. W., and White, A. P. (1981). The approximate randomization test as an alternative to the F test in analysis of variance. British Journal of Mathematics and Statistical Psychology, 34, 243-52.

Strodtbeck, F. L. (1958). Family interaction, values and achievement. In McClelland, D. C., (Ed.), Talent and society. New York: Van Nostrand.

Strumpfer, D. J. (1981). Towards a more socially responsive psychology. South African Journal of Psychology, Vol. 11(1), 18-28.

Suls, J. (1982). Social support, interpersonal relations and health: benefits and liabilities. In Sanders, G., and Suls, J. (Eds.), The social psychology of health and illness. Hillsdale, NJ: Erlbaum.

Thoits, P. A. (1982). Conceptual, methodological, and theoretical problems in studying social support as a buffer against life stress. Journal of Health and Social Behavior, June, Vol. 23, 145-159.

Tjosvold, D. (1982). Effects on approach to controversy on superiors' incorporation of subordinates' information in decision-making. Journal of Applied Psychology, 67, 189-193.

Tjosvold, D. (1984). Cooperation theory and organizations. Human relations. September, Vol. 37(9) 743-67.

Tjosvold, D. (1990). Power in cooperative and competitive organizational contexts. Journal of Social Psychology, April, Vol. 130(2) 249-58.

Tjosvold, D., and Johnson, D. W. (1978). Controversy within a cooperative or competitive context and cognitive perspective-taking. Contemporary Educational Psychology, 3, 376-386.

Tjosvold, D., Johnson, D. W., and Lerner, J. (1981). Effects of affirmation and acceptance on incorporation of opposing information in problem solving. Journal of Social Psychology, 114, 103-10.

Tjosvold, D., Johnson, D. W., and Johnson, R. T. (1984). Influence strategy, perspective-taking, and relationships between high- and low-power individuals in cooperative and competitive contexts. The Journal of Psychology, 116, 187-202.

Tracey, T. J., and Sedlacek, W. E. (1985). The relationship of noncognitive variables to academic success: a longitudinal comparison by race. Journal of College Student Personnel, September, Vol. 26(5), 405-410.

Triandis, H. C. (1972). The analysis of subjective culture. New York: Wiley Interscience.

Tucker, L.R. (1955). The objective definition of simple structure in linear factor analysis. Psychometrika, 20, 209-225.

Tukey, J. W. (1954). Causation, regression, and path analysis. In Kempthorne, O., Bancroft, T. A., Gowen, J. W., and Lush, J. L. (Eds.), Statistics and mathematics in biology (pp. 35-66). Ames, IA: Iowa State College Press.

Vaillant, G. (1974). Natural history of male psychological health: II. Some antecedents of healthy adult adjustment. Archives of General Psychiatry, 31, 15-22.

Vaillant, G. (1977). Adaptation to life. Boston, Little, Brown.

Valle, V. A., and Frieze, I. (1976). Stability of causal attributions as a mediator in changing expectations for success. Journal of Personality and Social Psychology, 33, 579-587.

Veit, C. T., and Ware, J. E. (1983). The structure of psychological distress and well-being in general populations. Journal of Consulting and Clinical Psychology, Vol. 51, No. 5, 730-42.

Vinokur, A., Schul, Y., and Caplan, R. D. (1987). Determinants of perceived social support: interpersonal transactions, personal outlook, and transient affective states. Journal of Personality and Social Psychology, Vol. 53, No. 6, 1137-1145.

Wack, J., and Rodin, J. (1978). Nursing homes for the aged: the human consequences of legislation-shaped environments. Journal of Social Issues, 34, 6-21.

Wallston, B. S., Alagna, S. W., DeVellis, B. M., and DeVellis, R. F. (1983). Social support and physical health. Health Psychology, 2(4), 367-391.

Wan, T. (1976). Predicting self-assessed health status: a multivariate approach. Health Services Research, 11, 646-477.

Ware, J. E., Johnsoton, S. A., Davis-Avery, A., and Brook, R. H. (1979). Conceptualization and measurement of health for adults in the health insurance study: Vol. III, mental health. Santa Monica, CA: Rand Corporation.

Ware, J. E., Davis-Avery, A., and Brook, R. H. (1980). Conceptualization and measurement of health for adults in the health insurance study: Vol. VI, analysis of relationships among health status measures. Santa Monica, CA: Rand Corporation.

Warr, P. (1982). Psychological aspects of employment and unemployment. Psychological Medicine, February, Vol. 12(1), 7-11.

Watson, T., Brown, M., and Swick, K. J. (1983). The relationship of parents' support to children's school achievement. Child Welfare, March-April, Vol. 62(2), 175-180.

Weiner, B. (1979). A theory of motivation for some classroom experiences. Journal of Educational Psychology, 71, 3-25.

Wiley, D. E. (1973). The identification problem for structural equation models with unmeasured variables. In Goldberger, A. S., and Duncan, O. D. (Eds.), Structural equation models in the social sciences. New York: Academic Press, pp. 69-83.

Williams, A. W., Ware, J. E., and Donald, C. A. (1981). A model of mental health, life events, and social supports applicable to general populations. Journal of Health and Social Behavior, 22, 324-336.

Wish, M., and Kaplan, S. J. (1974). Situational and relational aspects of interpersonal communication. Personality and Social Psychology Bulletin, 1974, Vol.1 (1), 234-236.

Wish, M., Deutsch, M., and Kaplan, S. J. (1976). Perceived dimensions of interpersonal relations. Journal of Personality and Social Psychology, 33, 409-20.

Wish, M., and Kaplan, S. J. (1977). Toward an implicit theory of interpersonal communication. Sociometry, 40, 234-46.

World Health Organization, (1948). Constitution, in basic documents, World Health Organization, Geneva.

Wortman, C. B., and Dunkel-Schetter, C. (1979). Interpersonal relationships and cancer: a theoretical analysis. Journal of Social Issues, 35, 120-155.

Wright, R. J., and DuCette, J. P. (1976). Locus of control and academic achievement in traditional and nontraditional settings, (cited in Lefcourt's study, 1982). Paper presented at American Education Research Association Convention, San, Francisco.

Wright, S. (1934). The method of path coefficients. Annals of Mathematical Statistics, 5, 161-215.

Zajonc, R., and Marin, I. (1967). Cooperation, competition, and interpersonal attitudes in small groups. Psychonomic Science, 7, 271-72.

Zhang, Q. (1990). Classroom climate patterns and cooperative learning. Unpublished paper, Teachers College, Columbia University.

Appendix A

Instruments

Intervention Scale I

Range 1-5

(strongly agree; agree; can't decide; disagree; strongly disagree)

Scoring direction: high = learned more or done better

How strongly do you agree with the following statements. Please CIRCLE the number that best describes your feelings.

a. During the past school year I have learned to work more effectively with other people in groups.

b. I have learned to put myself in the other person's shoes when I disagree with someone.

c. When I work in groups, I can expect the other group members to do my share of the work if I don't get around to doing it.

Intervention Scale II

Range 1-5

(no improvement; very little improvement; some improvement; much improvement; great improvement)

Scoring direction: high = more improvement

Please think about the way you handle arguments (conflicts), and about how you get along and work with others in groups. Please rate your improvement in the following areas.

a. handling conflicts or arguments with friends or peers

b. working with others in groups

c. handling conflicts/arguments with my family

If you have a job within the past year, please answer "d" & "e"

d. working with other employees on the job

e. handling conflicts/arguments on the job

Social Support Scale

Range 1-5

(very upsetting; upsetting; both upsetting and supportive; supportive; very supportive)

Scoring direction: high = more supportive

Please rate which areas are upsetting or helpful to you, when dealing with problems. Circle the one number on each line that best describes your feelings.

a. My family

b. My school

c. My work/job

d. My friends

e. My boyfriend/girlfriend

f. Myself

g. My religion or church

h. My neighborhood

i. Social workers, counselors

j. Police

k. Government agencies (unemployment, public assistance, ADC, AFDC, child care)

l. Health clinics, hospitals, doctors

School Disciplinary Issues Scale

Range 1-4

(everyday problem; frequent problem; not often a problem; not a problem)

Scoring direction: high = more of a problem

How much of a problem are the following disciplinary situations at your school?

a. Drugs used on campus
b. stealing
c. drinking liquor, beer, win on campus
d. weapons on campus
e. violence, student fighting or vandalism
f. cutting classes
g. students coming late to classes
h. students refusing to obey school policy
i. disrespect for teachers or staff
j. disrespect for other students

Victimization Scale

Range 1-6
(everyday; three times a week; once a week; once or twice a month; almost never; never)

Scoring direction: high = more frequently victimized

In the following statements about threats, attacks and theft, please think about both verbal insults and "put downs" as well as physical threats and attacks (fighting, shoving, weapons, etc.). Please think about things that have happened in the last three months. Circle the number that is closest to your answer. Remember to think only about things that have happened to you in the past three months.

a. I was insulted or threatened with words.
b. Someone forced me to hand over money or things (books, bus passes, jewelry) at school.
c. I received sexual attention from school or security staff that offended me.
d. I was affended that name calling or swearing.
e. I was afraid that someone would physically hurt me at school.
f. Students bring weapons to school to protect themselves.
g. I have been pressured or threatened to have sex with someone I did not want to have sex with.

h. Things have been stolen from my locker, lunch table, desk at school.

i. I have had things damaged by someone wanting to "get back at me".

j. I was physically attacked in school.

Self-Esteem Scale

Range 1-4

(strongly agree; agree; disagree; strongly disagree)

Scoring direction: high = high self-esteem

How do you feel about the following statements? Please circle the number that best describes your feelings.

a. I feel that I am a person of worth, equal to other people (R)

b. I feel that I have a number of good qualities (R)

c. All in all. I tend to think that I am a failure

d. I am able to do things as well as most other people (R)

e. I feel that I do not have much to be proud of

f. I have a positive attitude toward myself (R)

g. On the whole, I am satisfied with myself (R)

h. I wish I could have more respect for myself

i. I feel useless at times

j. At times I think that I am no good at all

Note:

(R) reversed scoring

Locus of Control Scale

Range 1-5

(strongly agree; agree; can't decide; disagree; strongly disagree)

Scoring direction: high = more internal orientation

How do you feel about the following statements? Please circle the number that best describes your feelings.

a. Good luck is more important than working hard •

b. Every time I get ahead, something or somebody stops me •

c. Planning is a waste of time, since it never works out anyway •

d. People who accept things the way they are are happier than people who try to make changes •

e. What happens to me is my own responsibility (R) i

f. When I make plans, I can usually make them work out successfully (R) i

Note:

•	external orientation
i	internal orientation
(R)	reversed scoring

General Well-Being Scale

Range 1-6 Depending on the type of question, the scaling of items somewhat differed, such as:

(excellent; very good; mostly good; both good and bad; mostly bad or low; bad, very low);
(extremely; very much; quite a bit; some; a little; not at all);
(every day; almost every day; about half the time; now and then; rarely; none of the time); etc.

Scoring direction: high=higher score, more of a construct

For clarity purposes, the GWB items are displayed as constituents of the appropriate scales. In the actual questionnaire the items were presented in a random fashion.

Positive Well-Being Subscale:

a. How have you been feeling in general? (R)
b. How happy, satisfied have you been with your personal life in the past month? (R)
c. Has life been full of interesting things during the past month? (R)
d. Have you felt cheerful during the past month? (R)

Self-Control Subscale:

a. Have you been in firm control of your behavior, thoughts and emotions or feelings during the past month? (R)
b. Have you had any reason to wonder if you were losing your mind, or losing control over the way you act, talk, think, feel, or remember things during the past month? (R)
c. Have you felt emotionally stable and sure of yourself during the past month? (R)

General Health Subscale:

a. How often were bothered by any illness, aches or pains during the past month?
b. Did you feel healthy enough to carry out the things you like to do or had to do in the past month? (R)
c. Have you been concerned, worried, or had any fears about your health during the past month?

Vitality Subscale:

a. How much energy or pep did you have or feel in the past month? (R)
b. Did you feel active or sluggish during the past month? (R)
c. Have you felt tired and worn out during the past month?
d. Have you been waking up feeling rested during the past month? (R)

Anxiety Subscale:

a. Have you been bothered by nervousness during the past month?
b. Have you been under or felt under any stress or pressure during the past month?
c. Have you been anxious, worried or upset during the past month?
d. Did you feel at ease and relaxed or stressed out and high-strung during the past month? (R)
e. Were you generally tense during the past month?

Depression Subscale:

a. Did you feel depressed during the past month?
b. Have you felt so sad, discouraged, hopeless, or had so many problems that you wondered if anything was worthwhile during the past month?
c. Have you felt depressed during the past month?

Note:

(R) reversed scoring

Appendix B

Additional Tables and Figures

Table 2 Distribution of Missing Values

Number of missing values	Number of cases
0	13
1	69
2	37
3	17
4	20
5	4
6	4
7	1
8	3
9	2
10	0
11	1
12	6
13	0
14	1
15	1
16	0
17	24
18	77
19	32
20	22
21	13
22	5
23	1
24	1
25	1
26	3
27	0
28	0
29	53
30	536
31	49
32	10
33	8
34	2
35	4
36	2
37	2
38	4
39	3
40	4
41	18
42	2
43	2
44	1
45	0
46	8
47	56

Table 3a

<u>T-tests between students who took both the pre- and post-test (Group 1) and students who took pre-test only (Group 2) for each site on the variables used in the analysis from the pre-test</u>

Site 1:

Variables	Group 1 Mean (SD) Cases	Group 2 Mean (SD) Cases	t	df	2-tail P
GDQUAL	3.44 (.57) 106	3.50 (.58) 252	-.85	356	.395
USELESS	2.62 (.90) 105	2.62 (.92) 245	-.05	348	.959
NOGOOD	2.93 (.93) 107	2.90 .99 245	.34	356	.735
LUCK	3.98 (.94) 107	4.02 (.96) 252	-.35	357	.726
WASTEOF	3.75 (.98) 106	3.69 (1.05) 252	.57	356	.568
PLANSUC	3.65 (.91) 107	3.78 (.98) 253	-1.13	358	.260
UPSET	3.14 (1.49) 102	3.37 (1.54) 241	-1.27	341	.206
TENSE	3.17 (1.32) 103	3.34 (1.55) 244	-1.00	345	.316
SAD	3.03 (1.70) 102	3.08 (1.81) 241	-.25	341	.799
FELTDEP	3.01 (1.44) 103	3.07 (1.50) 244	-.37	345	.713

Table 3a. (Continued)

Site 1:

Variables	Group 1 Mean (SD) Cases	Group 2 Mean (SD) Cases	t	df	2-tail P
CHEER	3.93 (1.42) 103	3.76 (1.53) 241	1.00	342	.317
WAKING	3.27 (1.37) 103	3.03 (1.53) 243	1.37	344	.171
LIFEFUL	3.53 (1.39) 103	3.45 (1.53) 242	.48	343	.634
STABLE	3.93 (1.39) 101	3.80 (1.53) 244	.72	343	.470
YOUPEP	4.38 (1.17) 100	4.36 (1.29) 244	.13	342	.897
YOUILL	4.31 (1.18) 104	4.31 (1.29) 242	.01	344	.990
WORRIED	4.28 (1.62) 102	3.97 (1.56) 242	1.69	342	.093

Table 3a. (Continued)

Site 2:

Variables	Group 1 Mean (SD) Cases	Group 2 Mean (SD) Cases	t	df	2-tail P
GDQUAL	3.49 (.56) 98	3.44 (.63) 212	.63	308	.531
USELESS	2.84 (.87) 94	2.77 (.89) 203	.65	295	.515
NOGOOD	3.01 (.94) 96	3.01 (.96) 211	.01	305	.994
LUCK	3.83 (1.11) 98	3.92 (.96) 216	-.73	312	.464
WASTEOF	3.81 (1.02) 97	3.56 (1.08) 212	1.94	307	.053
PLANSUC	3.67 (.98) 98	3.56 (1.03) 213	.89	309	.375
UPSET	2.84 (1.48) 95	3.23 (1.64) 208	-1.97	301	.050
TENSE	2.77 (1.51) 95	3.19 (1.65) 203	-2.09	296	.037
SAD	2.42 (1.53) 93	2.95 (1.73) 205	-2.52	296	.012
FELTDEP	2.71 (1.56) 93	3.06 (1.53) 204	-1.81	295	.071
CHEER	4.19 (1.61) 93	3.91 (1.49) 206	1.47	297	.142

Table 3a. (Continued)

Site 2:

Variables	Group 1 Mean (SD) Cases	Group 2 Mean (SD) Cases	t	df	2-tail P
WAKING	3.42 (1.75) 93	3.35 (1.55) 201	.35	292	.726
LIFEFUL	3.72 (1.67) 92	3.83 (1.43) 204	-.58	294	.559
STABLE	4.05 (1.63) 91	3.97 (1.50) 203	.46	292	.646
YOUPEP	4.60 (1.34) 94	4.40 (1.31) 207	1.22	299	.224
YOUILL	4.62 (1.41) 95	4.40 (1.34) 208	1.32	301	.189
WORRIED	4.57 (1.64) 94	4.17 (1.64) 205	1.98	297	.049

Table 3a. (Continued)

Site 3:

Variables	Group 1 Mean (SD) Cases	Group 2 Mean (SD) Cases	t	df	2-tail P
GDQUAL	3.47 (.67) 131	3.49 (.61) 221	-.40	350	.693
USELESS	2.78 (.91) 127	2.92 (.89) 213	-1.39	338	.164
NOGOOD	2.97 (.96) 131	3.13 (.93) 219	-1.57	348	.117
LUCK	4.08 (.83) 133	3.85 (1.01) 224	2.14	355	.033
WASTEOF	3.66 (1.12) 129	3.78 (1.00) 223	-1.01	350	.311
PLANSUC	3.69 (1.01) 133	3.75 (1.02) 223	-.51	354	.607
UPSET	2.94 (1.49) 130	2.84 (1.54) 215	.57	343	.568
TENSE	2.82 (1.43) 130	2.83 (1.59) 212	-.04	340	.967
SAD	2.68 (1.75) 130	2.46 (1.75) 212	1.14	340	.254
FELTDEP	2.71 (1.47) 130	2.65 (1.48) 212	.37	340	.709
CHEER	3.87 (1.35) 129	4.14 (1.44) 210	-1.75	337	.082

Table 3a. (Continued)

Site 3:

Variables	Group 1 Mean (SD) Cases	Group 2 Mean (SD) Cases	t	df	2-tail P
WAKING	3.33 (1.51) 125	3.37 (1.55) 210	-.22	333	.823
LIFEFUL	3.47 (1.54) 129	3.76 (1.54) 207	-1.72	334	.086
STABLE	3.86 (1.56) 128	4.04 (1.63) 208	-1.02	334	.307
YOUPEP	4.41 (1.24) 132	4.48 (1.35) 217	-.52	347	.604
YOUILL	4.29 (1.37) 133	4.42 (1.30) 217	-.86	348	.389
WORRIED	4.40 (1.75) 132	4.28 (1.65) 212	.66	342	.511

Table 3b

T-tests between students who attended school for three consecutive cycles (Group 1) and those who were absent for at least one cycle during the three consecutive cycles (Group 2) beginning from cycle 4, 5, and 6, respectively

Variables	Group 1 Mean (SD) Cases	Group 2 Mean (SD) Cases	t	df	2-tail P
Amark4[a]	71.67 (13.11) 120	70.69 (12.81) 177	.64	295	.525
Bmark4	71.25 (12.57) 88	73.43 (13.11) 149	-1.26	235	.210
Cmark4	70.03 (12.56) 88	71.68 (13.33) 147	-.94	233	.350
Dmark4	73.36 (13.43) 132	74.37 (13.31) 180	-.66	310	.511
Emark4	73.89 (12.94) 131	73.79 (13.34) 171	.06	300	.950
Fmark4	71.76 (13.24) 130	70.75 (13.40) 169	.65	297	.514
Amark5	75.46 (13.81) 169	72.94 (13.21) 203	1.79	370	.073
Bmark5	74.89 (12.49) 193	71.43 (12.76) 229	2.80	420	.005
Cmark5	74.46 (12.46) 111	71.59 (13.70) 195	1.82	304	.070

Note [a]This means a student's grade for Class A in cycle 4.

Table 3b. (Continued)

Variables	Group 1 Mean (SD) Cases	Group 2 Mean (SD) Cases	t	df	2-tail P
Dmark5	73.01 (12.31) 117	72.56 (12.94) 188	.30	303	.764
Emark5	74.51 (13.89) 190	70.24 (12.87) 203	3.16	391	.002
Fmark5	73.35 (13.79) 188	70.70 (13.24) 209	1.95	395	.052
Amark6	74.43 (14.27) 155	70.06 (12.98) 182	2.94	335	.004
Bmark6	77.11 (12.75) 151	74.39 (13.36) 148	1.8[illegible]	297	.073
Cmark6	72.69 (12.05) 77	76.21 (10.94) 98	-2.02	173	.044
Dmark6	75.85 (10.48) 86	78.97 (10.82) 96	-1.97	180	.050
Emark6	76.07 (13.57) 147	76.00 (12.90) 102	.04	247	.965
Fmark6	76.82 (13.25) 123	75.67 (12.90) 88	.63	209	.530

Table 3c

T-tests by site between students who attended school for three consecutive cycles (Group 1) and those who were absent for at least one cycle during the three consecutive cycles (Group 2) beginning from cycle 4, 5, and 6, respectively

Site 1:

Variables	Group 1 Mean (SD) Cases	Group 2 Mean (SD) Cases	t	df	2-tail P
Amark4	75.15 (14.34) 13	73.29 (11.89) 51	.48	62	.631
Bmark4	74.80 (13.18) 30	76.62 (12.69) 61	-.64	89	.526
Cmark4	74.47 (13.52) 30	75.08 (13.81) 59	-.20	87	.841
Dmark4	73.48 (14.20) 23	76.17 (13.43) 52	-.79	73	.434
Emark4	74.09 (12.74) 23	74.48 (11.64) 44	-.13	65	.900
Fmark4	75.48 (13.01) 21	76.69 (13.17) 39	-.34	58	.733

Table 3c. (Continued)

Site 2:

Variables	Group 1 Mean (SD) Cases	Group 2 Mean (SD) Cases	t	df	2-tail P
Amark4	79.06 (12.61) 50	76.71 (13.00) 41	.87	89	.385
Bmark4 (missing)					
Cmark4 (missing)					
Dmark4	76.50 (14.37) 54	80.02 (14.04) 43	-1.21	95	.229
Emark4	78.42 (12.70) 55	81.00 (12.51) 45	-1.02	98	.311
Fmark4	74.16 (13.95) 57	74.63 (14.86) 48	-.17	103	.869

Table 3c. (Continued)

Site 3:

Variables	Group 1 Mean (SD) Cases	Group 2 Mean (SD) Cases	t	df	2-tail P
Amark4	64.39 (8.73) 57	66.24 (11.71) 85	-1.02	140	.311
Bmark4	68.75 (11.54) 55	70.49 (12.61) 85	-.83	138	.409
Cmark4	66.98 (10.93) 55	68.61 (12.02) 85	-.81	138	.418
Dmark4	70.24 (11.53) 55	70.41 (11.63) 85	-.09	138	.930
Emark4	69.09 (11.73) 53	69.46 (12.98) 82	-.17	133	.867
Fmark4	67.63 (11.57) 52	65.64 (10.51) 82	1.03	132	.307

Table 3c. (Continued)

Site 1:

Variables	Group 1 Mean (SD) Cases	Group 2 Mean (SD) Cases	t	df	2-tail P
Amark5	74.18 (12.38) 62	76.26 (12.64) 40	-.83	100	.409
Bmark5	73.77 (12.04) 78	72.25 (13.50) 59	.69	135	.490
Cmark5	75.53 (12.36) 76	77.80 (14.54) 56	-.97	130	.334
Dmark5	74.81 (12.23) 78	78.06 (11.40) 48	-1.49	124	.139
Emark5	74.31 (12.79) 67	78.03 (12.80) 37	-1.42	102	.159
Fmark5	72.54 (12.79) 63	74.77 (13.68) 43	-.86	104	.394

Table 3c. (Continued)

Site 2:

Variables	Group 1 Mean (SD) Cases	Group 2 Mean (SD) Cases	t	df	2-tail P
Amark5	77.71 (15.61) 70	75.63 (15.31) 22	.55	90	.586
Bmark5	76.78 (12.68) 79	75.21 (11.84) 29	.58	106	.561
Cmark5 (missing)					
Dmark5 (missing)					
Emark5	76.80 (14.91) 88	72.20 (14.25) 35	1.56	121	.121
Fmark5	74.51 (14.58) 91	73.64 (14.53) 36	.30	125	.763

Table 3c. (Continued)

Site 3:

Variables	Group 1 Mean (SD) Cases	Group 2 Mean (SD) Cases	t	df	2-tail P
Amark5	73.35 (12.08) 37	71.57 (12.87) 141	.76	176	.451
Bmark5	73.17 (12.78) 36	70.31 (12.54) 141	1.22	175	.225
Cmark5	72.14 (12.52) 35	69.09 (12.55) 139	1.29	172	.199
Dmark5	68.94 (11.65) 36	70.76 (12.94) 139	-.76	173	.447
Emark5	69.11 (11.92) 35	67.52 (11.53) 131	.72	164	.471
Fmark5	71.76 (13.55) 34	68.55 (12.32) 130	1.33	162	.186

Table 3c. (Continued)

Site 1:

Variables	Group 1 Mean (SD) Cases	Group 2 Mean (SD) Cases	t	df	2-tail P
Amark6	72.56 (12.81) 52	74.80 (14.37) 35	-.76	85	.448
Bmark6	74.33 (12.19) 48	78.21 (12.99) 34	-1.38	80	.172
Cmark6	71.67 (13.28) 52	74.74 (13.75) 38	-1.07	88	.290
Dmark6	75.66 (11.28) 58	79.89 (12.09) 35	-1.71	91	.092
Emark6	75.44 (13.13) 52	77.26 (13.24) 31	-.61	81	.545
Fmark6	70.26 (12.63) 39	73.36 (13.93) 22	-.89	59	.378

Table 3c. (Continued)

Site 2:

Variables	Group 1 Mean (SD) Cases	Group 2 Mean (SD) Cases	t	df	2-tail P
Amark6	77.97 (15.67) 70	72.35 (15.40) 34	1.72	102	.088
Bmark6	80.27 (12.91) 77	77.60 (13.29) 25	.89	100	.374
Cmark6 (missing)					
Dmark6 (missing)					
Emark6	76.73 (14.77) 75	70.86 (15.02) 29	1.81	102	.073
Fmark6	81.31 (12.69) 64	70.83 (14.50) 24	3.32	86	.001

Table 3c. (Continued)

Site 3:

Variables	Group 1 Mean (SD) Cases	Group 2 Mean (SD) Cases	t	df	2-tail P
Amark6	69.85 (11.61) 33	67.90 (11.24) 113	.87	144	.386
Bmark6	72.88 (11.21) 26	72.03 (13.14) 89	.30	113	.765
Cmark6	74.80 (8.84) 25	77.15 (8.71) 60	-1.13	83	.262
Dmark6	76.25 (8.78) 28	78.44 (10.09) 61	-.99	87	.325
Emark6	75.25 (9.93) 20	78.62 (10.34) 42	-1.21	60	.229
Fmark6	75.25 (10.70) 20	79.64 (10.15) 42	-1.57	60	.123

Table 7

Squared Multiple Correlation Coefficients--Reliability Indices[a] (Model I)

Y-variables	R-square	X-variables	R-square
Pfamup	.328	Gdqual	.226
Pschup	.666	Useless	.646
Pwrkup	.355	Nogood	.676
Pfriup	.331	Luck	.144
Pgdqua	.306	Wasteof	.392
Pusele	.681	Plansuc	.293
Pnogoo	.668	Upset	.505
Pluck	.338	Tense	.502
Pwastp	.557	Sad	.583
Pplans	.122	Feltdep	.606
Achieve	.994	Cheer	.649
Pupset	.466	Waking	.303
Ptense	.599	Lifeful	.452
Psad	.534	Stable	.590
Pfeltd	.561	Youpep	.187
Pcheer	.715	Youill	.269
Pwakin	.400	Worried	.517
Plifef	.456	Peffgr	.656
Pstabl	.549	Pothdo	.008
Pyoupe	.195	Pothsh	.433
Pyouil	.284	Pimpgp	.708
Pworri	.645	Pimpcr	.607
		Pimpem	.647
		Pimpct	.491
		Pimpjb	.542

Note

[a]The squared multiple correlation coefficients are the lower bounds of reliability for each of the variables. Since the specific components of reliability are not included. In psychometrics, they are R-squares between the items and the true scores.

Table 8

Power Analyses for Ten Parameters[a] (Model I)

Parameters	H1:	H0:	Alpha[b]	Df	Lambda[c]	Power
Gamma(1,7)	.25	0	.05	1	2.78	.39
Beta(5,1)	-.28	0	.05	1	1.46	.23
Beta(6,1)	.33	0	.05	1	2.35	.33
Beta(5,2)	-.40	0	.05	1	3.68	.48
Beta(6,2)	.29	0	.05	1	1.90	.28
Beta(2,4)	.02	0	.05	1	2.42	.34
Beta(6,4)	.03	0	.05	1	4.06	.52
Gamma(1,3)	-.17	0	.05	1	1.87	.28
Gamma(2,2)	.11	0	.05	1	.15	.07
Gamma(6,4)	.29	0	.05	1	4.17	.53

Note

[a]The power analysis is conducted for each of the ten parameters under the assumption that H1, the alternative hypothesis, is true, while H0 is being tested. The values of power are obtained from the program of Lispower by Joreskog and sorbom (1989).

[b]Alpha is the significance level for each test.

[c]Lambda is the non-centrality parameter for each test.

Table 12

Squared Multiple Crrelation Coefficients--Reliability

Indices (Model II)

Y-variables	R-square	X-variables	R-square
Pdrugs	.636	Gdqua	.215
Pdrink	.703	Useless	.604
Pweapo	.634	Nogood	.719
Pviole	.403	Luck	.154
Pjewel	.809	Wasteof	.391
Phurt	.572	Plansuc	.286
Pphys	.739	Upset	.471
Pgdqua	.277	Tense	.476
Pusele	.652	Sad	.596
Pnogoo	.675	Feltdep	.642
Pluck	.358	Peffgr	.606
Pwastp	.495	Pothdo	.008
Pplans	.099	Pothsh	.479
Pupset	.448	Pimpgp	.700
Ptense	.537	Pimpem	.641
Psad	.543	Pimpcr	.595
Pfeltd	.590	Pimpct	.490
Pcheer	.657	Pimpjb	.537
Plifef	.477		
Pstabl	.566		
Pwakin	.418		

Table 13

ML Estimates of Model II (Unstandardized)

Coefficients		ML Estimates	Standard Error	T-value
LAMBDA Y				
	Lambda 1,1	1.000	.000	.000
	Lambda 2,1	1.055	.101	10.441
	Lambda 3,1	.998	.100	9.965
	Lambda 4,1	.796	.103	7.737
	Lambda 5,2	1.189	.110	10.832
	Lambda 6,2	1.000	.000	.000
	Lambda 7,2	1.136	.107	10.595
	Lambda 8,3	-.101	.129	-.783
	Lambda 9,3	1.000	.000	.000
	Lambda 10,3	1.017	.133	7.653
	Lambda 8,4	.933	.214	4.361
	Lambda 11,4	1.000	.000	.000
	Lambda 12,4	1.192	.194	6.159
	Lambda 13,4	.526	.165	3.195
	Lambda 14,5	.871	.113	7.707
	Lambda 15,5	.955	.113	8.431
	Lambda 16,5	.960	.113	8.476
	Lambda 17,5	1.000	.000	.000
	Lambda 18,5	-.177	.089	-1.984
	Lambda 18,6	1.108	.144	7.673
	Lambda 19,6	1.000	.000	.000
	Lambda 20,6	1.089	.142	7.673
	Lambda 21,6	.936	.138	6.801
LAMBDA X				
	Lambda 1,1	-.041	.147	-.275
	Lambda 2,1	1.000	.000	.000
	Lambda 3,1	1.091	.156	7.010
	Lambda 1,2	.765	.216	3.538
	Lambda 4,2	.628	.166	3.783
	Lambda 5,2	1.000	.000	.000
	Lambda 6,2	.855	.177	4.819
	Lambda 7,3	.856	.104	8.269
	Lambda 8,3	.861	.103	8.323
	Lambda 9,3	.964	.103	9.352
	Lambda 10,3	1.000	.000	.000
	Lambda 11,4	1.000	.000	.000
	Lambda 12,4	-.114	.123	-.930
	Lambda 13,4	.889	.194	4.585
	Lambda 14,5	1.085	.103	10.486
	Lambda 15,5	1.039	.104	10.014
	Lambda 16,5	1.000	.000	.000
	Lambda 17,5	.908	.105	8.614
	Lambda 18,5	.950	.105	9.067

Coefficients		ML Estimates	Standard Error	T-value
BETA				
	Beta 2,1	.089	.089	.995
	Beta 3,2	-.089	.106	-.834
	Beta 4,2	-.224	.089	-2.505
	Beta 5,2	.187	.093	2.009
	Beta 4,3	.142	.108	1.314
	Beta 5,3	-.369	.106	-3.473
	Beta 6,3	.278	.105	2.639
	Beta 4,5	.008	.126	.061
GAMMA				
	Gamma 1,1	.028	.136	.207
	Gamma 2,1	-.119	.124	-.967
	Gamma 3,1	.375	.140	2.683
	Gamma 4,1	-.093	.118	-.786
	Gamma 5,1	.179	.125	1.426
	Gamma 6,1	-.081	.124	-.651
	Gamma 1,2	-.359	.207	-1.735
	Gamma 2,2	-.196	.196	-1.000
	Gamma 3,2	.113	.208	.543
	Gamma 4,2	.554	.192	2.890
	Gamma 5,2	-.067	.176	-.381
	Gamma 6,2	.317	.177	1.796
	Gamma 1,3	.043	.108	.397
	Gamma 2,3	-.147	.098	-1.491
	Gamma 3,3	.025	.108	.234
	Gamma 4,3	-.051	.121	-.418
	Gamma 5,3	.611	.109	5.590
	Gamma 6,3	-.093	.094	-.995
	Gamma 1,4	-.059	.126	-.471
	Gamma 2,4	.039	.112	.344
	Gamma 3,4	-.016	.122	-.130
	Gamma 4,4	.109	.097	1.118
	Gamma 5,4	.111	.107	1.040
	Gamma 6,4	.107	.108	.992
	Gamma 1,5	-.013	.114	-.113
	Gamma 2,5	-.290	.106	-2.735
	Gamma 3,5	.239	.118	2.018
	Gamma 4,5	-.029	.095	-.301
	Gamma 5,5	.158	.104	1.523
	Gamma 6,5	-.094	.101	-.927
PSI				
	Psi 1,1	.572	.106	5.375
	Psi 2,2	.465	.092	5.080
	Psi 3,3	.467	.100	4.672
	Psi 4,4	.119	.052	2.271
	Psi 5,5	.276	.068	4.064
	Psi 6,6	.353	.085	4.161

(Continue to the next page)

Coefficients	ML Estimates	Standard Error	T-value
THETA EPS (TE)			
TE 1,1	.364	.058	6.232
TE 2,2	.292	.055	5.308
TE 3,3	.366	.059	6.260
TE 4,4	.597	.077	7.762
TE 5,5	.191	.048	3.943
TE 6,6	.428	.059	7.229
TE 7,7	.261	.050	5.239
TE 8,8	.723	.098	7.382
TE 9,9	.348	.082	4.246
TE 10,10	.325	.083	3.922
TE 11,11	.642	.090	7.114
TE 12,12	.519	.000	.000
TE 13,13	.901	.108	8.365
TE 14,14	.552	.075	7.332
TE 15,15	.463	.069	6.741
TE 16,16	.457	.068	6.692
TE 17,17	.410	.065	6.281
TE 18,18	.342	.063	5.427
TE 19,19	.523	.075	6.977
TE 20,20	.434	.070	6.210
TE 21,21	.582	.079	7.348
THETA DELTA (TD)			
TD 1,1	.785	.103	7.601
TD 2,2	.396	.087	4.547
TD 3,3	.281	.094	2.990
TD 4,4	.846	.105	8.062
TD 5,5	.609	.095	6.403
TD 6,6	.714	.098	7.304
TD 7,7	.529	.072	7.305
TD 8,8	.524	.072	7.273
TD 9,9	.404	.063	6.379
TD 10,10	.358	.061	5.893
TD 11,11	.394	.130	3.021
TD 12,12	.992	.115	8.635
TD 13,13	.521	.114	4.582
TD 14,14	.300	.049	6.087
TD 15,15	.359	.054	6.683
TD 16,16	.405	.058	7.047
TD 17,17	.510	.067	7.613
TD 18,18	.463	.063	7.392

Table 14 Total Effects for Model II

Total effects of ETA on ETA

	Pcrime	Pvictim	Pesteem	Pcontrol	Pnmentl	Ppmentl
Pcrime	.000	.000	.000	.000	.000	.000
Pvictim	.089	.000	.000	.000	.000	.000
Pesteem	-.008	-.089	.000	.000	.000	.000
Pcontrol	-.021	-.235	.140	.000	.008	.000
Pnmentl	.019	.219	-.369	.000	.000	.000
Ppmentl	-.002	-.025	.278	.000	.000	.000

Standard errors for the total effects of ETA on ETA

	Pcrime	Pvictim	Pesteem	Pcontrol	Pnmentl	Ppmentl
Pcrime	.000	.000	.000	.000	.000	.000
Pvictim	.089	.000	.000	.000	.000	.000
Pesteem	.012	.106	.000	.000	.000	.000
Pcontrol	.023	.087	.093	.000	.126	.000
Pnmentl	.021	.097	.106	.000	.000	.000
Ppmentl	.004	.031	.105	.000	.000	.000

Total effects of KSI on ETA

	Resteem	Rcontrol	Rnmentl	Group	Conflict
Pcrime	.028	-.359	.043	-.059	-.013
Pvictim	-.117	-.228	-.143	.033	-.291
Pesteem	.386	.133	.038	-.019	.264
Pcontrol	-.012	.623	-.009	.100	.074
Pnmentl	.014	-.159	.570	.124	.006
Ppmentl	.027	.354	-.083	.102	-.020

Standard errors for the total effects of KSI on ETA

	Resteem	Rcontrol	Rnmentl	Group	Conflict
Pcrime	.136	.207	.108	.126	.114
Pvictim	.124	.191	.099	.113	.107
Pesteem	.140	.206	.107	.123	.115
Pcontrol	.113	.203	.085	.101	.091
Pnmentl	.121	.184	.110	.113	.101
Ppmentl	.118	.183	.096	.111	.099

Table 15

ML Estimates of Variances and Covariances of Exogenous Variables for Model II

Coefficients (PHI)	ML Estimates (Unstandardized)	ML Estimates (Standardized)	Standard Error	T-value
Phi 2,1	.247	.509	.070	3.521
Phi 3,1	-.272	-.437	.072	-3.778
Phi 4,1	.118	.195	.068	1.747
Phi 5,1	.092	.154	.059	1.557
Phi 3,2	-.154	-.308	.062	-2.498
Phi 4,2	.193	.396	.067	2.862
Phi 5,2	.194	.402	.061	3.190
Phi 4,3	-.104	-.167	.067	-1.547
Phi 5,3	-.103	-.167	.060	-1.726
Phi 5,4	.233	.388	.069	3.370
Phi 1,1	.604		.129	4.667
Phi 2,2	.391		.112	3.502
Phi 3,3	.642		.117	5.503
Phi 4,4	.606		.162	3.745
Phi 5,5	.595		.111	5.371

ML Estimates of Variances of Endogenous Variables (Unstandardized)

Variables	ML Estimates
Pcrime	.636
Pvictim	.572
Pesteem	.652
Pcontrol	.358
Pnmentl	.590
Ppmentl	.477

Table 17

ML Estimates of the Factor Structures--Site 1

(Unstandardized)

Coefficients	ML Estimates	ndard Error	T-value
LAMBDA X			
Lambda 1,1	1.00	.00	.00
Lambda 2,1	.94	.21	4.54
Lambda 3,1	.72	.19	3.78
Lambda 4,2	.95	.27	3.47
Lambda 5,2	1.00	.00	.00
Lambda 6,2	.60	.19	3.18
Lambda 7,3	1.18	.19	6.26
Lambda 8,3	1.00	.00	.00
Lambda 9,3	1.01	.16	6.30
Lambda 10,4	1.00	.00	.00
Lambda 11,4	-.15	.19	-.78
Lambda 12,4	.92	.24	3.81
Lambda 13,5	1.11	.20	5.50
Lambda 14,5	1.35	.24	5.59
Lambda 15,5	1.00	.00	.00
Lambda 16,5	.78	.20	3.90
THETA DELTA (TD)			
Pdrugs	.06	.06	1.09
Pweapo	.25	.06	3.89
Pviole	.39	.07	5.35
Pfamup	.92	.20	4.62
Pschup	.20	.13	1.52
Pwrkup	.56	.11	5.14
Pjewel	.33	.09	3.49
Phurt	.42	.09	4.61
Pphys	.22	.07	3.29
Peffgr	.30	.16	1.91
Pothdo	1.21	.21	5.86
Pothsh	.42	.14	2.94
Pimpgp	.34	.09	3.69
Pimpcr	.38	.12	3.14
Pimpct	.67	.13	5.05
Pimpjb	.83	.15	5.47
TAU X (TX)			
TX 1	1.28	.08	15.38
TX 2	1.39	.07	18.52
TX 3	1.53	.07	21.28
TX 4	3.53	.13	27.87
TX 5	3.80	.10	36.46
TX 6	3.45	.10	36.02
TX 7	1.39	.12	11.52
TX 8	1.32	.10	13.82

Coefficients	ML Estimates	Standard Error	T-value
TX 9	1.30	.11	11.30
TX 10	4.08	.09	43.70
TX 11	3.27	.11	31.05
TX 12	3.87	.10	39.80
TX 13	3.55	.12	28.83
TX 14	3.30	.13	24.59
TX 15	3.47	.12	27.85
TX 16	3.42	.12	29.72
KAPPA (KA)			
KA 1	.09	.10	.88
KA 2	-.26	.14	-1.84
KA 3	-.02	.14	-.11
KA 4	-.29	.14	-2.06
KA 5	.00	.14	.02

Table 18

ML Estimates of Factor Variances and Covariances for Site 1 (Unstandardized)

Coefficients (PHI)	ML Estimates	Standard Error	T-value
Phi 2,1	.05	.06	.83
Phi 3,1	.08	.06	1.51
Phi 4,1	-.10	.06	-1.54
Phi 5,1	.02	.06	.27
Phi 3,2	-.12	.08	-1.48
Phi 4,2	.07	.09	.79
Phi 5,2	.19	.09	2.20
Phi 4,3	-.21	.09	-2.34
Phi 5,3	-.19	.08	-2.22
Phi 5,4	.20	.09	2.14
Phi 1,1	.29	.08	3.63
Phi 2,2	.51	.17	2.97
Phi 3,3	.51	.15	3.36
Phi 4,4	.59	.20	2.93
Phi 5,5	.54	.18	2.92

Assessment of fit for Site 1:

Goodness of fit index: .846
Root mean square residual: .078

Table 19

ML Estimates of Factor Structures--Site 2 (Unstandardized)

Coefficients	ML Estimates	Standard Error	T-value
LAMBDA X			
Lambda 1,1	1.00	.00	.00
Lambda 2,1	1.00	.13	7.87
Lambda 3,1	.70	.14	5.10
Lambda 4,2	.82	.20	4.05
Lambda 5,2	1.00	.00	.00
Lambda 6,2	.42	.19	2.22
Lambda 7,3	1.45	.14	10.57
Lambda 8,3	1.00	.00	.00
Lambda 9,3	1.37	.15	8.84
Lambda 10,4	1.00	.00	.00
Lambda 11,4	-.08	.29	-.29
Lambda 12,4	.78	.23	3.32
Lambda 13,5	1.04	.16	6.60
Lambda 14,5	.83	.15	5.64
Lambda 15,5	1.00	.00	.00
Lambda 16,5	.84	.16	5.19
THETA DELTA (TD)			
TD 1,1	.11	.04	2.89
TD 2,2	.08	.03	2.21
TD 3,3	.32	.06	5.50
TD 4,4	1.16	.22	5.35
TD 5,5	.20	.00	.00
TD 6,6	1.11	.19	5.73
TD 7,7	.09	.06	1.66
TD 8,8	.17	.04	4.39
TD 9,9	.40	.08	4.78
TD 10,10	.29	.12	2.52
TD 11,11	1.62	.28	5.87
TD 12,12	.59	.12	4.97
TD 13,13	.43	.12	3.58
TD 14,14	.65	.13	4.91
TD 15,15	.66	.15	4.49
TD 16,16	.89	.17	5.17
TAU X (TX)			
TX 1	1.28	.08	16.38
TX 2	1.39	.07	18.52
TX 3	1.53	.07	21.28
TX 4	3.53	.13	27.87
TX 5	3.80	.10	36.46
TX 6	3.45	.10	36.02
TX 7	1.39	.12	11.52
TX 8	1.32	.10	13.82
TX 9	1.30	.11	11.30
TX 10	4.08	.09	43.70

Coefficients	ML Estimates	Standard Error	T-value
TX 11	3.27	.11	31.05
TX 12	3.87	.10	39.80
TX 13	3.55	.12	28.83
TX 14	3.30	.13	24.59
TX 15	3.47	.12	27.85
TX 16	3.42	.12	29.72
KAPPA (KA)			
KA 1	.00		
KA 2	.00		
KA 3	.00		
KA 4	.00		
KA 5	.00		

Table 20

ML Estimates of Factor Variances and Covariances for Site 2 (Unstandardized)

Coefficients (PHI)	ML Estimates	Standard Error	T-value
Phi 2,1	-.24	.07	-3.20
Phi 3,1	.18	.06	2.98
Phi 4,1	-.08	.06	-1.41
Phi 5,1	-.14	.08	-1.73
Phi 3,2	-.19	.08	-2.39
Phi 4,2	.18	.09	2.05
Phi 5,2	.21	.12	1.80
Phi 4,3	-.10	.07	-1.46
Phi 5,3	-.27	.10	-2.66
Phi 5,4	.45	.13	3.60
Phi 1,1	.34	.08	4.20
Phi 2,2	.60	.14	4.41
Phi 3,3	.47	.11	4.35
Phi 4,4	.39	.15	2.65
Phi 5,5	.92	.26	3.51

Assessment of fit for Site 2:

Goodness of fit index: .829
Root mean square residual: .085

Table 21

ML Estimates of the Factor Structures--Site 3

(Unstandardized)

Coefficients	ML Estimates	Standard Error	T-value
LAMBDA X			
Lambda 1,1	1.00	.00	.00
Lambda 2,1	.96	.05	18.10
Lambda 3,1	1.08	.06	17.48
Lambda 4,2	1.00	.04	25.32
Lambda 5,2	1.00	.00	.00
Lambda 6,2	.94	.03	34.47
Lambda 7,3	1.05	.06	16.24
Lambda 8,3	1.00	.00	.00
Lambda 9,3	.97	.06	16.14
Lambda 10,4	1.00	.00	.00
Lambda 11,4	.80	.04	19.17
Lambda 12,4	.92	.03	30.40
Lambda 13,5	1.08	.04	28.69
Lambda 14,5	1.03	.04	26.12
Lambda 15,5	1.00	.00	.00
Lambda 16,5	1.03	.05	22.18
THETA DELTA (TD)			
TD 1,1	.43	.09	4.94
TD 2,2	.17	.05	3.60
TD 3,3	.27	.07	4.05
TD 4,4	1.17	.22	5.36
TD 5,5	.15	.07	2.11
TD 6,6	.49	.11	4.66
TD 7,7	.18	.06	3.14
TD 8,8	.50	.10	5.05
TD 9,9	.16	.05	3.27
TD 10,10	.46	.12	3.95
TD 11,11	1.58	.28	5.56
TD 12,12	.59	.13	4.66
TD 13,13	.19	.06	3.01
TD 14,14	.36	.08	4.47
TD 15,15	.71	.14	5.23
TD 16,16	.79	.15	5.27
KAPPA (KA)			
KA 1	1.65	.10	16.16
KA 2	3.45	.11	30.76
KA 3	1.38	.13	10.82
KA 4	3.91	.11	36.65
KA 5	3.11	.15	21.37

Table 22

ML Estimates of Factor Variances and Covariances for Site 3 (Unstandardized)

Coefficients (PHI)	ML Estimates	Standard Error	T-value
Phi 2,1	-.05	.07	-.71
Phi 3,1	.03	.07	.39
Phi 4,1	-.06	.06	-1.12
Phi 5,1	-.03	.07	-.40
Phi 3,2	-.17	.10	-1.67
Phi 4,2	.14	.08	1.67
Phi 5,2	.31	.11	2.83
Phi 4,3	-.01	.08	-.08
Phi 5,3	-.25	.11	-2.31
Phi 5,4	.20	.09	2.19
Phi 1,1	.31	.07	4.30
Phi 2,2	.71	.15	4.92
Phi 3,3	.73	.15	4.74
Phi 4,4	.33	.11	3.08
Phi 5,5	.80	.16	5.06

Assessment of fit for Site 3:

Goodness of fit index: .850
Root mean square residual: .103

Overall Goodness of Fit across Three Sites:

Chi-square with 305 df ....347.53 P=.047

Table 23

ML Estimates for Mean Structures

1. Unstandardized factor loadings

Variables	Crime	Support	Factors Victim	Group	Conflict
Pdrugs	1.00	.00	.00	.00	.00
Pweapo	.95	.00	.00	.00	.00
Pviole	.75	.00	.00	.00	.00
Pfamup	.00	.87	.00	.00	.00
Pschup	.00	1.00	.00	.00	.00
Pwrkup	.00	.67	.00	.00	.00
Pjewel	.00	.00	1.28	.00	.00
Phurt	.00	.00	1.00	.00	.00
Pphys	.00	.00	1.16	.00	.00
Peffgr	.00	.00	.00	1.00	.00
Pothdo	.00	.00	.00	-.18	.00
Pothsh	.00	.00	.00	.81	.00
Pimpgp	.00	.00	.00	.00	1.06
Pimpcr	.00	.00	.00	.00	1.04
Pimpct	.00	.00	.00	.00	1.00
Pimpjb	.00	.00	.00	.00	.87

2. Unstandardized Error Variances and Intercepts

Variables	Theta Delta	Tau X
Pdrugs	.18	1.37
Pweapo	.16	1.37
Pviole	.34	1.55
Pfamup	1.09	3.59
Pschup	.20	3.78
Pwrkup	.74	3.44
Pjewel	.20	1.43
Phurt	.38	1.40
Pphys	.26	1.36
Peffgr	.27	4.09
Pothdo	1.40	3.23
Pothsh	.53	3.82
Pimpgp	.32	3.53
Pimpcr	.49	3.31
Pimpct	.67	3.40
Pimpjb	.86	3.38

3. Factor Variance-Covariance Matrix

(Site 2)			Factors		
	Crime	Support	Victim	Group	Conflict
Crime	.32				
Support	-.22	.56			
Victim	.20	-.18	.59		
Group	-.08	.17	-.09	.40	
Conflict	-.13	.18	-.29	.41	.84

(Site 1)			Factors		
	Crime	Support	Victim	Group	Conflict
Crime	.28				
Support	.04	.51			
Victim	.08	-.10	.45		
Group	-.08	.08	-.20	.63	
Conflict	-.02	.21	-.19	.22	.63

(Site 3)			Factors		
	Crime	Support	Victim	Group	Conflict
Crime	.41				
Support	-.06	.76			
Victim	.02	-.16	.53		
Group	-.10	.20	-.01	.60	
Conflict	-.06	.35	-.20	.23	.80

4. Mean Structure Matrix

Sites	Crime	Support	Victim	Group	Conflict
		Factor Means			
Site 1	.11	-.25	-.07	-.30	.02
Site 2	.00	.00	.00	.00	.00
Site 3	.28	-.31	.00	-.21	-.17

5. Scaled Mean Structure Matrix[a]

Sites	Crime	Support	Victim	Group	Conflict
		Factor Means			
Site 1	-.02	-.063	-.047	-.13	.07
Site 2	-.13	.187	.023	.17	.05
Site 3	.15	-.123	.023	-.04	-.12

Note

[a]The scaled factor means have been computed such that the weighted mean (We have set the sample sizes equal across groups, thus the weight here is a unity.) over the groups is zero for each factor.

Table 24

Assessment of Fit for the Mean Structures Model

1. Assessment of fit for Site 2

Goodness of fit index:	.776
Root mean square residual:	.106

2. Assessment of fit for Site 1

Goodness of fit index:	.829
Root mean square residual:	.091

3. Assessment of fit for Site 3

Goodness of fit index:	.833
Root mean square residual:	.085

4. Overall Assessment of Goodness of Fit

Chi-square with 359 df	442.92	P=.002

Figure 6 Equivalent Models with three Y-variables

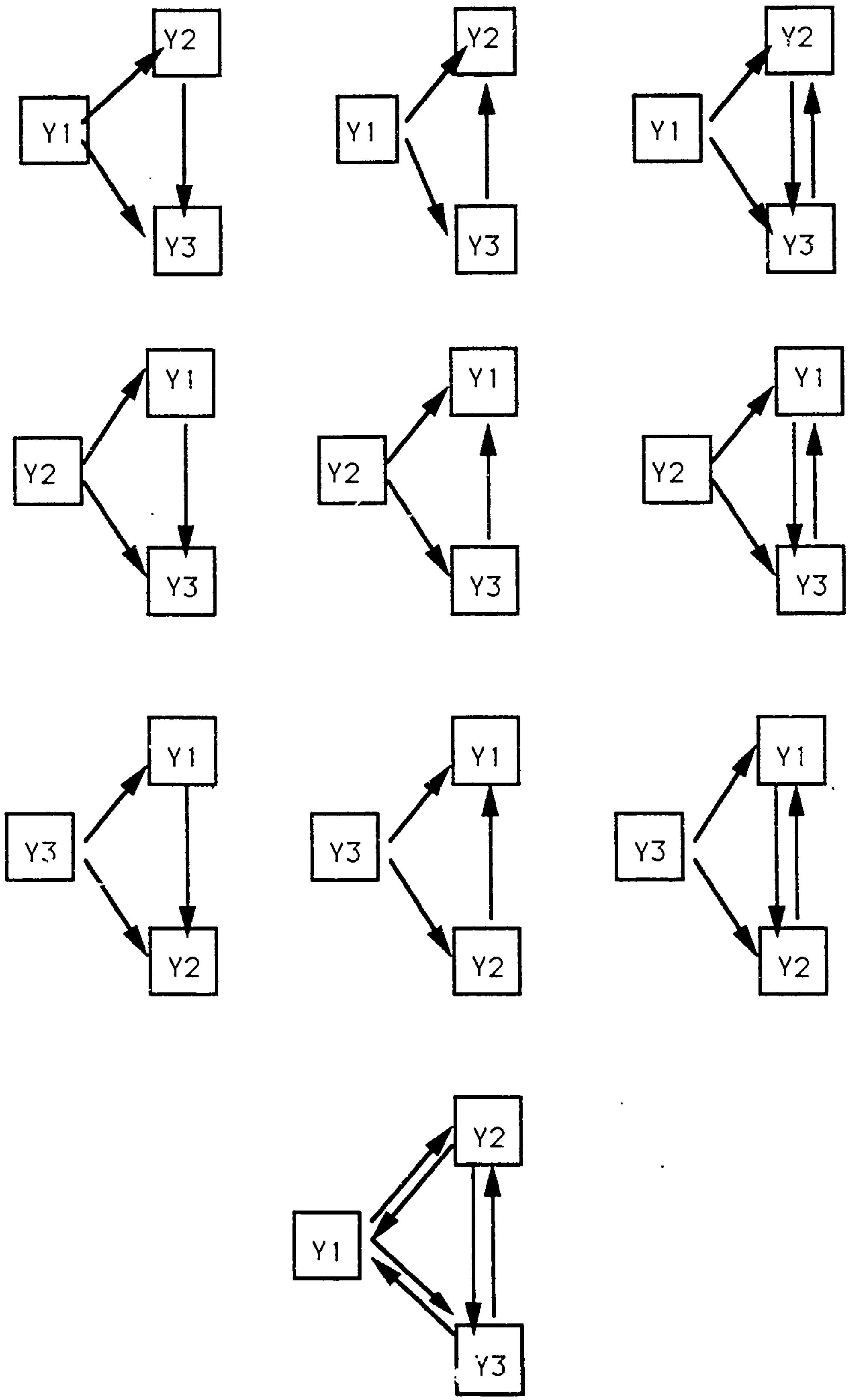
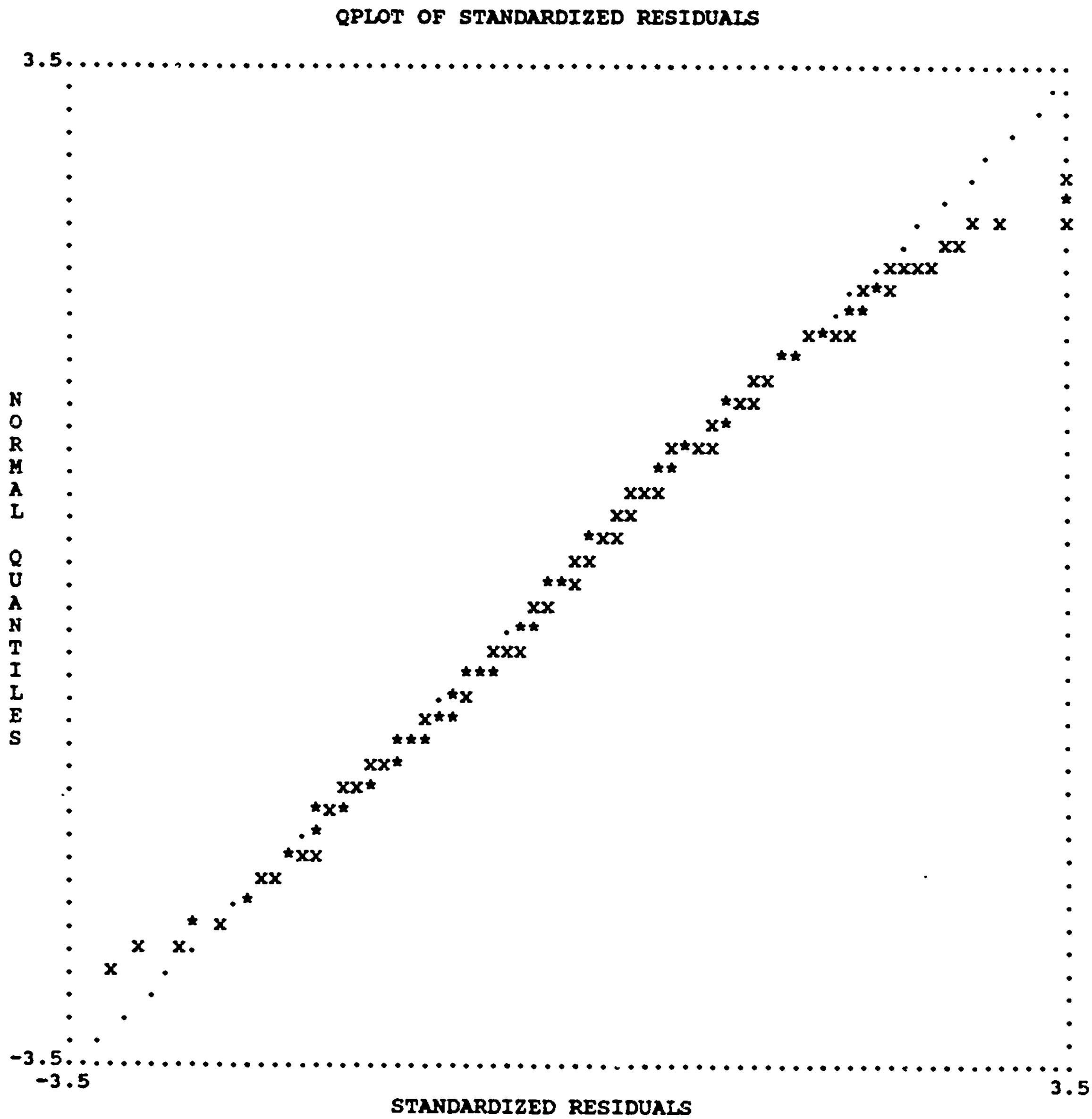


Figure 9 Q-Plot of Model II



Appendix C

Description of the Intervention Plans

Cooperative learning training

The model of cooperative learning that was employed in our training was the one developed by David and Roger Johnson of the Center for Cooperative Learning at the University of Minnesoda. Our trainers were trained by the Jognsons and they made available to the teachers of AHS various books and teaching materials developed by the Johnsons to assist the teachers in implementing cooperative learning in their classrooms.

According to the Johnsons model, there are four major elements that must be included for small group learning to be truely cooperative. These four elements are: (a) positive interdependence Students must perceive that they "sink or swim together." This may be achieved through mutual goals (goal interdependence), division of labor (task interdependence), dividing materials, resources, or information among group members (resource interdependence), assigning students roles (role interdependence), and by giving joint rewards (reward interdependence). (b) face-to-face interaction among students (c) individual accountability Each group member is responsible for the overall group performance. (d) interpersonal and small group skills.

Our goals for the teachers being trained in cooperative learning were to (1) have them laerning a model of cooperative learning; (2) have them learning how to create cooperative lessons for use in their classrooms; (3) motivate them to use cooperative learning in their classes; (4) have them practice the cooperative laerning in their classes; and (5) have them help one another in developing their knowledge and skills relating to cooperative learning.

The cooperative learning training took place at an initial two-day workshop in August of 1988 and through periodic after-school workshops during the school year, combined with individual conferences once a week with teachers who requested them. We note that usually it takes about three years of training and experience in implementing the Johnson and Johnson model before teachers feel that they are highly proficient in its use. A typical pattern of development is that teachers starting experimenting on a trial basis with cooperative learning in some of their classes during the first year of training; in the second year they become more committed to cooperative learning and use it regularly but they do not yet feel fully comfortable with it; in the third year they start to be comfortable and innovative in using it in their classes.

Conflict resolution training

The training in conflict resolution was designed for two different but interactive audiances: students and the faculty. The training content has been descibed in Chapter three of this thesis. Since conflict resolution does not require a systematic group practice, training for individual students was the main focus of the intervention.

The faculty was responsible for student training, which took place in Family Groups (a form of class), occupational education classes, and orientation classes. These classes meet three times a week. Family Group is a chance for students to informally discuss personal problems they encountered at school or in their lives with their peers and their family group faculty advisor. Topics vary from class to class depending on what the students bring up. The faculty also present topics and activities in Family Group to meet particular group needs. Orientation classes provide incoming students with behavioral and academic guidelines.

Faculty received support from our staff in implementing the training through role-playing, discussion groups, group projects, etc. There was a suggested series of topics , however, each teacher was free to move at her/his own pace. The topics are listed below. These were tentative and could be changed if necessary to meet the emerging needs of the students and staff.

1. Violence prevention From a previous study conducted at one of the sites, it was learned that many students at the implementing school use fighting and verbal abuse as ways to resolve some of the conflicts they find themselves in. While fighting at school is an immediate cause for dismissal, there appears to be a culture clash between "the rules at school" and the expected norms of "the rest."

The Boston Violence Prevention Project (Prothow-Smith, 1987) was designed to show teens from communities where violence is the "street norms" how "at risk" they are of injury or death. The curriculum points out that many of the homocides which occur result from a fight with someone you know where a weapon and/or alcohol is present. The curriculum also helps students to understand that anger is normal, but that they have a choice as to how they can express their anger. The final chapter teach violence prevention strategies.

It was assumed that the discussion flowing from this topic would help students and teachers to distinguish those situations that were negotiable (conflict between people who know each other and have reasons to maintain that relationship) from situations that were non-negotiable (random and instrumental violence) and which occur often in the "street." These conflicts are best avoided if possible; negotiation would not be appropriate.

2. Basic negotiation skills Deutsch has outlined a theoretical framework for conflict resolution in the schools, and Raider has developed a method for teaching negotiation skills. With these as a guide, and borrowing from the seventy-six lesson plans created by the San Francisco Community Borad, faculty and students were to explore the topics of values, needs, climate, and listening and communication skills.

3. Applying negotiation skills Role play, group

activities, and discussion groups were to be utilized in the practical application of negotiation skills to students' lives in home, school, and work settings.

4. Basic mediation skills Mediation skills were to be tought. The students would have the opportunity to practice mediation by facilitating constructive conflict resolution. It was assumed that helping others would reinforce one's desire to use the newly learned skills in one's own conflicts.

Appendix D

Computer Programs for LISREL Models

File 1: Program for Model I

```
GET FILE='ACAD:[Q_Z_M_DEUTS]ALL.SYS'
RECODE PUPSET PTENSE PSAD PFELTD PLIFEF
       PCHEER PSTABL PWAKIN PYOUPE
       UPSET TENSE SAD FELTDEP LIFEFUL
       CHEER STABLE WAKING YOUPEP
       (1=6) (2=5) (3=4) (4=3) (5=2) (6=1)
RECODE PDRUGS PDRINK PWEAPO PVIOLE DRUGS
       DRINK WEAPON VIOLEN
       (1=4) (2=3) (3=2) (4=1)
RECODE PJEWEL PHURT PLOCKE PPHYS JEWEL HURT LOCKER PHYS
       (1=6) (2=5) (3=4) (4=3) (5=2) (6=1)
RECODE PEFFGR POTHSH PCNFLC PLESSA PSUMMA
       (1=5) (2=4) (4=2) (5=1)
RECODE PSIMPR PWAPRO PBRPRO PSLPRO PFIPRO PMIRPR PEATPR
       (5=1) (4=2) (2=4) (1=5)
RECODE PWORTH PGDQUA PAMABL PPOSAT PSATSE
       (1=4) (2=3) (3=2) (4=1)
RECODE WORTH GDQUAL AMABLE POSATT SATSELF
       (1=4) (2=3) (3=2) (4=1)
RECODE PHAPPE PPLANS HAPPENS PLANSUC PEASYH POFFNT PCOUNT
       (1=5) (2=4) (4=2) (5=1)
COMPUTE ESTEEM1=MEAN(PWORTH, PGDQUA, PIFAIL)
COMPUTE ESTEEM2=MEAN(PAMABL, PMUCH)
COMPUTE ESTEEM3=MEAN(PUSELE, PNOGOO)
COMPUTE ESTEEM4=MEAN(PPOSAT, PSATSE, PWISHM)
COMPUTE ESTEEM11=MEAN(WORTH, GDQUAL, IFAIL)
COMPUTE ESTEEM12=MEAN(AMABLE, MUCH)
COMPUTE ESTEEM13=MEAN(USELESS, NOGOOD)
COMPUTE ESTEEM14=MEAN(POSATT, SATSELF, WISHMO)
COMPUTE LOCUS1=MEAN(PSTOPS, PWASTP, PHAPPI)
COMPUTE LOCUS2=MEAN(PHAPPE, PPLANS)
COMPUTE LOCUS11=MEAN(STOPSME, WASTEOF, HAPPIER)
COMPUTE LOCUS12=MEAN(HAPPENS, PLANSUC)
COMPUTE PGROUP=MEAN(PEFFGR, PIMPGP)
COMPUTE ACHIEVE=MEAN.2(MATH,READ,HIST,SCNC,GLOB)

PRELIS
  /VARIABLES=PFAMUP PSCHUP PWRKUP PFRIUP
    PGDQUA PUSELE PNOGOO PLUCK PWASTP PPLANS ACHIEVE
    PUPSET PTENSE PSAD PFELTD PCHEER PWAKIN PLIFEF PSTABL
    PYOUPE PYOUIL PWORRI
    GDQUAL USELESS NOGOOD LUCK WASTEOF PLANSUC
    UPSET TENSE SAD FELTDEP CHEER WAKING LIFEFUL STABLE
    YOUPEP YOUILL WORRIED
    PEFFGR POTHDO POTHSH PIMPGP PIMPCR PIMPEM PIMPCT PIMPJB
  /MISSING = PAIRWISE
  /TYPE=COV

LISREL
 /LATENT MODEL: MODEL1 WITH SOCIAL SUPPORT
 /SIZE=1024
 /DA  NI = 47
```

```
/LA
/'PFAMUP' 'PSCHUP' 'PWRKUP' 'PFRIUP'
/ PGDQUA PUSELE PNOGOO PLUCK PWASTP PPLANS
/'ACHIEVE'
/'PUPSET' 'PTENSE' 'PSAD' 'PFELTD'
/'PCHEER' 'PWAKIN' 'PLIFEF' 'PSTABL'
/'PYOUPE' 'PYOUIL' 'PWORRI'
/ GDQUAL USELESS NOGOOD LUCK WASTEOF PLANSUC
/'UPSET' 'TENSE' 'SAD' 'FELTDEP'
/'CHEER' 'WAKING' 'LIFEFUL' 'STABLE'
/'YOUPEP' 'YOUILL' 'WORRIED'
/ PEFFGR POTHDO POTHSH PIMPGP PIMPCR PIMPEM PIMPCT PIMPJB
/MO NY=22 NE=7 NX=25 NK=7 BE=FU,FI GA=FU,FI PS=DI
/LE
/'PSUPPORT' 'PESTEEM' 'PLOCUS'
/'ACHIEVE' 'PNMENTL' 'PPMENTL'
/'PPHYSIC'
/LK
/'RESTEEM' 'RCONTROL'
/'RNMENTL' 'RPMENTL'
/'RPHYSIC' 'GROUP' 'CONF'
/FR  LY(1,1) LY(3,1) LY(4,1) LY(5,2) LY(7,2) LY 5,3 LY(8,3)
   LY(10,3) LY(13,5) LY(14,5) LY(15,5) LY(16,6) LY(17,6)
   LY(19,6) LY(20,6) LY(20,7) LY(22,7)
   LX(1,1) LX(3,1) LX(1,2) LX(4,2) LX(6,2) LX(8,3) LX(9,3)
   LX(10,3) LX(11,4) LX(12,4) LX(14,4) LX(15,4) LX(15,5)
   LX(17,5) LX(19,6) LX(20,6) LX(21,7) LX(22,7)
   LX(24,7) LX(25,7)
/VA  1 LY 2 1 LY 6 2 LY 9 3 LY 11 4 LY 12 5 LY 18 6 LY 21 7
     LX 2 1 LX 5 2 LX 7 3 LX 13 4 LX 16 5 LX 18 6 LX 23 7
/FI  TE 11
/VA  .5 TE 11
/FR  GA(2,1) GA(1,2) GA(2,2) GA(3,2) GA(5,3) GA(1,4)
     GA(6,4) GA(7,5) GA(1,6) GA(2,6) GA(1,7) GA(5,7)
     GA(3,6) GA(6,7) GA(7,7) GA(1,3)
/FR  BE(2,1) BE(3,1) BE(7,1) BE(5,1) BE(6,1) BE(3,2)
     BE(6,4) BE(4,3) BE(2,4) BE(3,5) BE(3,7) BE(5,2)
     BE(6,2)  BE(7,5)
/OU  SE TV EF RE RS SS SC MI TO
```

File 2: Program for Model II

```
GET FILE=ALL.SYS
RECODE PUPSET PTENSE PSAD PFELTD PLIFEF
       PCHEER PSTABL PWAKIN PYOUPE
       UPSET TENSE SAD FELTDEP LIFEFUL
       CHEER STABLE WAKING YOUPEP
       (1=6) (2=5) (3=4) (4=3) (5=2) (6=1)
RECODE PDRUGS PDRINK PWEAPO PVIOLE
       DRUGS DRINK WEAPON VIOLEN
       (1=4) (2=3) (3=2) (4=1)
RECODE PJEWEL PHURT PLOCKE PPHYS JEWEL HURT LOCKER PHYS
       (1=6) (2=5) (3=4) (4=3) (5=2) (6=1)
RECODE PEFFGR POTHSH PCNFLC PLESSA PSUMMA
       (1=5) (2=4) (4=2) (5=1)
RECODE PSIMPR PWAPRO PBRPRO PSLPRO PFIPRO PMIRPR PEATPR
       (5=1) (4=2) (2=4) (1=5)
RECODE PWORTH PGDQUA PAMABL PPOSAT PSATSE
       (1=4) (2=3) (3=2) (4=1)
RECODE WORTH GDQUAL AMABLE POSATT SATSELF
       (1=4) (2=3) (3=2) (4=1)
RECODE PHAPPE PPLANS HAPPENS PLANSUC PEASYH POFFNT PCOUNT
       (1=5) (2=4) (4=2) (5=1)
COMPUTE ESTEEM1=MEAN(PWORTH, PGDQUA, PIFAIL)
COMPUTE ESTEEM2=MEAN(PAMABL, PMUCH)
COMPUTE ESTEEM3=MEAN(PUSELE, PNOGOO)
COMPUTE ESTEEM4=MEAN(PPOSAT, PSATSE, PWISHM)
COMPUTE ESTEEM11=MEAN(WORTH, GDQUAL, IFAIL)
COMPUTE ESTEEM12=MEAN(AMABLE, MUCH)
COMPUTE ESTEEM13=MEAN(USELESS, NOGOOD)
COMPUTE ESTEEM14=MEAN(POSATT, SATSELF, WISHMO)
COMPUTE LOCUS1=MEAN(PSTOPS, PWASTP, PHAPPI)
COMPUTE LOCUS2=MEAN(PHAPPE, PPLANS)
COMPUTE LOCUS11=MEAN(STOPSME, WASTEOF, HAPPIER)
COMPUTE LOCUS12=MEAN(HAPPENS, PLANSUC)
COMPUTE PGROUP=MEAN(PEFFGR, PIMPGP)
COMPUTE ACHIEVE=MEAN.2(MATH,READ,HIST,SCNC,GLOB)
RECODE PEFFGR (6=5)

PRELIS
  /VARIABLES= PDRUGS PDRINK PWEAPO PVIOLE PJEWEL PHURT PPHYS
    PGDQUA PUSELE PNOGOO PLUCK PWASTP PPLANS
    PUPSET PTENSE PSAD PFELTD PCHEER PLIFEF PSTABL PWAKIN
    GDQUAL USELESS NOGOOD LUCK WASTEOF PLANSUC
    UPSET TENSE SAD FELTDEP
    PEFFGR POTHDO POTHSH PIMPGP PIMPEM PIMPCR PIMPCT PIMPJB
  /MATRIX=OUT(*)
  /MISSING = PAIRWISE
  /TYPE=CORRELATION
LISREL
 /"STRUCTURAL MODEL WITH SCHOOL CRIME & VICTIMIZATION"
 /MATRIX=IN(*)
 /DA  NI = 39  NO=151
 /LA
```

```
/'PDRUGS' 'PDRINK' 'PWEAPO' 'PVIOLE' 'PJEWEL' 'PHURT'
 'PPHYS'
 'PGDQUA' 'PUSELE' 'PNOGOO' 'PLUCK' 'PWASTP' 'PPLANS'
 'PUPSET' 'PTENSE' 'PSAD' 'PFELTD'
 'PCHEER' 'PLIFEF' 'PSTABL' 'PWAKIN'
 'GDQUAL' 'USELESS' 'NOGOOD' 'LUCK' 'WASTEOF' 'PLANSUC'
 'UPSET' 'TENSE' 'SAD' 'FELTDEP'
 'PEFFGR' 'POTHDO' 'POTHSH' 'PIMPGP' 'PIMPEM'
 'PIMPCR' 'PIMPCT' 'PIMPJB'
/MO  NY=21 NE=6 NX=18 NK=5   BE=FU BE=FI GA=FU GA=FR PS=DI
/LE
/'PCRIME' 'PVICTIM' 'PESTEEM' 'PCONTROL' 'PNMENTL'
 'PPMENTL'
/LK
/'RESTEEM' 'RCONTROL' 'RNMENTL' 'COOP' 'CONF'
/FR  LY(2,1) LY(3,1) LY(4,1) LY(5,2) LY(7,2) LY(8,3)
   LY(8,4) LY(10,3) LY(12,4) LY(13,4) LY(14,5) LY(15,5)
   LY(16,5) LY(18,5) LY(18,6) LY(20,6) LY(21,6) LX(1,1)
   LX(3,1) LX(1,2) LX(4,2) LX(6,2) LX(7,3) LX(8,3) LX(9,3)
   LX(12,4) LX(13,4) LX(14,5) LX(15,5) LX(17,5) LX(18,5)
/VA  1 LY(1,1) LY(6,2) LY(9,3) LY(11,4) LY(17,5) LY(19,6)
   LX(2,1) LX(5,2) LX(10,3) LX(11,4) LX(16,5)
/FI  TE(12,12)
/VA  .519 TE(12,12)
/FR  BE(2,1) BE(3,2) BE(4,3) BE(4,2) BE(5,2) BE(5,3)
     BE(6,3) BE(4,5)
/OU  SE TV EF RS RE MI SS SC TO
```

File 3: Program for Model III

```
MATRIX DATA FILE=GROUP2.MTX
  /VARIABLES=X1 X2 X3 X4 X5 X6 X7 X8 X9 X10 X11 X12
      X13 X14 X15 X16 X17 X18 X19 X20
  /FORMAT=FREE FULL
  /CONTENTS=MEAN SD CORR
  /SPLIT=SPL

LISREL
 /TEST OF GROUP MEAN DIFFERENCES: SITE 1
 /SIZE=1024
 /DA  NI=20  NO=70 NG=3
 /ME  FI=GROUP2.MTX; SD  FI=GROUP2.MTX; KM FU FI=GROUP2.MTX
 /LA
 /'PDRUGS' 'PDRINK' 'PWEAPO' 'PVIOLE'
 /'PFAMUP' 'PSCHUP' 'PWRKUP' 'PFRIUP'
 /'PJEWEL' 'PHURT' 'PLOCKE' 'PPHYS'
 /'PEFFGR' 'POTHDO' 'POTHSH'
 /'PIMPGP' 'PIMPCR' 'PIMPCT' 'PIMPEM' 'PIMPJB'
 /SE
 /1 3 4 5 6 7 9 10 12 13 14 15 16 17 18 20/
 /MO  NX=16 NK=5  TX=FR KA=FR
 /LK
 /'CRIME' 'SUPPORT' 'VICTIM' 'GROUP' 'CONFLICT'
 /FR  LX(2,1) LX(3,1) LX(4,2) LX(6,2) LX(7,3) LX(9,3)
   LX(11,4) LX(12,4) LX(13,5) LX(14,5) LX(16,5)
 /VA  1 LX(1,1) LX(5,2) LX(8,3) LX(10,4) LX(15,5)
 /OU  SE TV RS MI SS  ND=2 TO

 /TEST OF GROUP MEAN DIFFERENCES: SITE 2
 /DA  NO=70
 /ME  FI=GROUP2.MTX; SD  FI=GROUP2.MTX; KM FU FI=GROUP2.MTX
 /LA
 /'PDRUGS' 'PDRINK' 'PWEAPO' 'PVIOLE'
 /'PFAMUP' 'PSCHUP' 'PWRKUP' 'PFRIUP'
 /'PJEWEL' 'PHURT' 'PLOCKE' 'PPHYS'
 /'PEFFGR' 'POTHDO' 'POTHSH'
 /'PIMPGP' 'PIMPCR' 'PIMPCT' 'PIMPEM' 'PIMPJB'
 /SE
  /1 3 4 5 6 7 9 10 12 13 14 15 16 17 18 20/
 /MO   LX=PS KA=FI TX=IN
 /LK
 /'CRIME' 'SUPPORT' 'VICTIM'  'GROUP' 'CONFLICT'
 /FI  TD 5,5
 /VA  .2 TD(5,5)
 /OU

 /TEST OF GROUP MEAN DIFFERENCES: SITE 3
 /DA  NO=70
 /ME  FI=GROUP2.MTX; SD  FI=GROUP2.MTX; KM FU FI=GROUP2.MTX
 /LA
 /'PDRUGS' 'PDRINK' 'PWEAPO' 'PVIOLE'
 /'PFAMUP' 'PSCHUP' 'PWRKUP' 'PFRIUP'
```

```
/'PJEWEL' 'PHURT' 'PLOCKE' 'PPHYS'
/'PEFFGR' 'POTHDO' 'POTHSH'
/'PIMPGP' 'PIMPCR' 'PIMPCT' 'PIMPEM' 'PIMPJB'
/SE
/1 3 4 5 6 7 9 10 12 13 14 15 16 17 18 20/
/MO  LX=PS KA=FR TX=FI
/LK
/'CRIME' 'SUPPORT' 'VICTIM' 'GROUP' 'CONFLICT'
/OU
```

File 4: Program for Model IV

```
MATRIX DATA FILE=GROUP20.MTX
  /VARIABLES=X1 X2 X3 X4 X5 X6 X7 X8 X9 X10 X11 X12 X13 X14
     X15 X16 X17 X18 X19 X20
  /FORMAT=FREE FULL
  /CONTENTS=MEAN SD CORR
  /SPLIT=SPL

LISREL
 /TEST OF GROUP MEAN DIFFERENCES: SITE 2
 /SIZE=1024
 /DA  NI=20  NO=70 NG=3
 /ME  FI=GROUP20.MTX; SD  FI=GROUP20.MTX; KM FU
      FI=GROUP20.MTX
 /LA
 /'PDRUGS' 'PDRINK' 'PWEAPO' 'PVIOLE'
 /'PFAMUP' 'PSCHUP' 'PWRKUP' 'PFRIUP'
 /'PJEWEL' 'PHURT' 'PLOCKE' 'PPHYS'
 /'PEFFGR' 'POTHDO' 'POTHSH'
 /'PIMPGP' 'PIMPCR' 'PIMPCT' 'PIMPEM' 'PIMPJB'
 /SE
 /1 3 4 5 6 7 9 10 12 13 14 15 16 17 18 20/
 /MO  NX=16 NK=5  TX=FR KA=FI
 /LK
 /'CRIME' 'SUPPORT' 'VICTIM' 'GROUP' 'CONFLICT'
 /FR  LX(2,1) LX(3,1) LX(4,2) LX(6,2) LX(7,3) LX(9,3)
   LX(11,4) LX(12,4) LX(13,5) LX(14,5) LX(16,5)
 /VA  1 LX(1,1) LX(5,2) LX(8,3) LX(10,4) LX(15,5)
 /FI  TD 5,5
 /VA  .2 TD(5,5)
 /OU  SE TV RS MI SS  ND=2 TO

 /TEST OF GROUP MEAN DIFFERENCES: SITE 1
 /DA  NO=70
 /ME  FI=GROUP20.MTX; SD  FI=GROUP20.MTX; KM FU
      FI=GROUP20.MTX
 /LA
 /'PDRUGS' 'PDRINK' 'PWEAPO' 'PVIOLE'
 /'PFAMUP' 'PSCHUP' 'PWRKUP' 'PFRIUP'
 /'PJEWEL' 'PHURT' 'PLOCKE' 'PPHYS'
 /'PEFFGR' 'POTHDO' 'POTHSH'
 /'PIMPGP' 'PIMPCR' 'PIMPCT' 'PIMPEM' 'PIMPJB'
 /SE
  /1 3 4 5 6 7 9 10 12 13 14 15 16 17 18 20/
 /MO   LX=IN KA=FR TX=IN TD=IN
 /LK
 /'CRIME' 'SUPPORT' 'VICTIM'  'GROUP' 'CONFLICT'
 /OU

 /TEST OF GROUP MEAN DIFFERENCES: SITE 3
 /DA  NO=70
 /ME  FI=GROUP20.MTX; SD  FI=GROUP20.MTX; KM FU
      FI=GROUP20.MTX
```

```
/LA
/'PDRUGS' 'PDRINK' 'PWEAPO' 'PVIOLE'
/'PFAMUP' 'PSCHUP' 'PWRKUP' 'PFRIUP'
/'PJEWEL' 'PHURT' 'PLOCKE' 'PPHYS'
/'PEFFGR' 'POTHDO' 'POTHSH'
/'PIMPGP' 'PIMPCR' 'PIMPCT' 'PIMPEM' 'PIMPJB'
/SE
/1 3 4 5 6 7 9 10 12 13 14 15 16 17 18 20/
/MO  LX=IN KA=FR TX=IN TD=IN
/LK
/'CRIME' 'SUPPORT' 'VICTIM' 'GROUP' 'CONFLICT'
/OU
```

File 5: Program for producing Sigma matrix[a]

```
GET FILE='ACAD:[Q_Z_M_DEUTS]ALL.SYS'
RECODE PUPSET PTENSE PSAD PFELTD PLIFEF PCHEER PSTABL PWAKIN
       PYOUE UPSET TENSE SAD FELTDEP LIFEFUL CHEER STABLE
       WAKING YOUPEP
       (1=6) (2=5) (3=4) (4=3) (5=2) (6=1)
RECODE PDRUGS PDRINK PWEAPO PVIOLE DRUGS DRINK WEAPON VIOLEN
       (1=4) (2=3) (3=2) (4=1)
RECODE PJEWEL PHURT PLOCKE PPHYS JEWEL HURT LOCKER PHYS
       (1=6) (2=5) (3=4) (4=3) (5=2) (6=1)
RECODE PEFFGR POTHSH PCNFLC PLESSA PSUMMA
       (1=5) (2=4) (4=2) (5=1)
RECODE PSIMPR PWAPRO PBRPRO PSLPRO PFIPRO PMIRPR PEATPR
       (5=1) (4=2) (2=4) (1=5)
RECODE PWORTH PGDQUA PAMABL PPOSAT PSATSE
       (1=4) (2=3) (3=2) (4=1)
RECODE WORTH GDQUAL AMABLE POSATT SATSELF
       (1=4) (2=3) (3=2) (4=1)
RECODE PHAPPE PPLANS HAPPENS PLANSUC PEASYH POFFNT PCOUNT
       (1=5) (2=4) (4=2) (5=1)
COMPUTE ESTEEM1=MEAN(PWORTH, PGDQUA, PIFAIL)
COMPUTE ESTEEM2=MEAN(PAMABL, PMUCH)
COMPUTE ESTEEM3=MEAN(PUSELE, PNOGOO)
COMPUTE ESTEEM4=MEAN(PPOSAT, PSATSE, PWISHM)
COMPUTE ESTEEM11=MEAN(WORTH, GDQUAL, IFAIL)
COMPUTE ESTEEM12=MEAN(AMABLE, MUCH)
COMPUTE ESTEEM13=MEAN(USELESS, NOGOOD)
COMPUTE ESTEEM14=MEAN(POSATT, SATSELF, WISHMO)
COMPUTE LOCUS1=MEAN(PSTOPS, PWASTP, PHAPPI)
COMPUTE LOCUS2=MEAN(PHAPPE, PPLANS)
COMPUTE LOCUS11=MEAN(STOPSME, WASTEOF, HAPPIER)
COMPUTE LOCUS12=MEAN(HAPPENS, PLANSUC)
COMPUTE PGROUP=MEAN(PEFFGR, PIMPGP)
COMPUTE ACHIEVE=MEAN.2(MATH,READ,HIST,SCNC,GLOB)

PRELIS
  /VARIABLES=PFAMUP PSCHUP PWRKUP PFRIUP
    PGDQUA PUSELE PNOGOO PLUCK PWASTP PPLANS ACHIEVE
    PUPSET PTENSE PSAD PFELTD PCHEER PWAKIN PLIFEF PSTABL
    PYOUPE PYOUIL PWORRI
    GDQUAL USELESS NOGOOD LUCK WASTEOF PLANSUC
    UPSET TENSE SAD FELTDEP CHEER WAKING LIFEFUL STABLE
    YOUPEP YOUILL WORRIED
    PEFFGR POTHDO POTHSH PIMPGP PIMPCR PIMPEM PIMPCT PIMPJB
  /MISSING = PAIRWISE
  /TYPE=COV

LISREL
 /COMPUTING COVARIANCE MATRIX FOR POWER ANALYSIS
 /SIZE=1024
 /DA  NI = 47
 /LA
 /'PFAMUP' 'PSCHUP' 'PWRKUP' 'PFRIUP'
```

```
/  PGDQUA PUSELE PNOGOO PLUCK PWASTP PPLANS
/'ACHIEVE'
/'PUPSET' 'PTENSE' 'PSAD' 'PFELTD'
/'PCHEER' 'PWAKIN' 'PLIFEF' 'PSTABL'
/'PYOUPE' 'PYOUIL' 'PWORRI'
/GDQUAL USELESS NOGOOD LUCK WASTEOF PLANSUC
/'UPSET' 'TENSE' 'SAD' 'FELTDEP' 'CHEER' 'WAKING' 'LIFEFUL'
 'STABLE'
/'YOUPEP' 'YOUILL' 'WORRIED'
/PEFFGR POTHDO POTHSH PIMPGP PIMPCR PIMPEM PIMPCT PIMPJB
/MO  NY=22  NE=7  NX=25  NK=7  BE=FU BE=FI GA=FU GA=FI
   PS=FI,DI TE=FI TD=FI PH=SD,FI
/LE
/'PSUPPORT' 'PESTEEM' 'PLOCUS'
/'ACHIEVE' 'PNMENTL' 'PPMENTL'
/'PPHYSIC'
/LK
/'RESTEEM' 'RCONTROL'
/'RNMENTL' 'RPMENTL'
/'RPHYSIC' 'GROUP' 'CONF'
/MA LY
/ .7 0 0 0 0 0 0
/  1 0 0 0 0 0 0
/ .7 0 0 0 0 0 0
/ .7 0 0 0 0 0 0
/  0 -.07 .7 0 0 0 0 0
/  0 1 0 0 0 0 0
/  0 1 0 0 0 0 0
/  0 0 .75 0 0 0 0
/  0 0 1 0 0 0 0
/  0 0 .45 0 0 0 0
/  0 0 0 1 0 0 0
/  0 0 0 0 1 0 0
/  0 0 0 0 1 0 0
/  0 0 0 0 1 0 0
/  0 0 0 0 1 0 0
/  0 0 0 0 0 1.3 0
/  0 0 0 0 0 .94 0
/  0 0 0 0 0 1 0
/  0 0 0 0 0 1 0
/  0 0 0 0 0 .6 .25
/  0 0 0 0 0 0 1
/  0 0 0 0 0 0 1.5
/MA LX
/  0 .7 0 0 0 0 0
/  1 0 0 0 0 0 0
/  1 0 0 0 0 0 0
/  0 .6 0 0 0 0 0
/  0 1 0 0 0 0 0
/  0 .87 0 0 0 0 0
/  0 0 1 0 0 0 0
/  0 0 1 0 0 0 0
/  0 0 1 0 0 0 0
/  0 0 1 0 0 0 0
```

```
/  0 0 0 1.2 0 0 0
/  0 0 0 .8 0 0 0
/  0 0 0 1 0 0 0
/  0 0 0 1.2 0 0 0
/  0 0 0 .4 .54 0 0
/  0 0 0 0 1 0 0
/  0 0 0 0 1.4 0 0
/  0 0 0 0 0 1 0
/  0 0 0 0 0 -.1 0
/  0 0 0 0 0 .8 0
/  0 0 0 0 0 0 1
/  0 0 0 0 0 0 .96
/  0 0 0 0 0 0 1
/  0 0 0 0 0 0 .9
/  0 0 0 0 0 0 .9
/MA BE
/  0 0 0 0 0 0 0
/  .31 0 0 .02 0 0 0
/  .11 .13 0 0 .03 0 .1
/  0 0 4.13 0 0 0 0
/  -.28 -.4 0 0 0 0 0
/  .33 .29 0 .03 0 0 0
/  .03 0 0 0 -.49 0 0
/MA GA
/  0 -.02 -.17 .03 0 .1 .25
/  .34 .11 0 0 0 0 0
/  0 .65 0 0 0 .14 0
/  0 0 0 0 0 0 0
/  0 0 .4 0 0 0 .15
/  0 0 0 .29 0 0 -.07
/  0 0 0 0 -.02 0 .01
/MA PH
/  .53 .23 .4 -.34 -.20 1.23 .24 .27 -.42 1.05
/  .15 .14 -.6 .19 .46 .11 .2 -.12 .18 .07 .54
/  .12 .25 -.16 .25 .07 .28 .92
/MA PS
/  .5 .4 .3 71  .6 .6 .2
/MA TE
/  1 .3 .7 .6 .3 .3 .3 .7 .5 .9 .5 1.2 1 1 1 .6 1.3 1.2 1.1
/  1.4 1.2 .9
/MA TD
/ .3 .3 .3 .9 .6 .7 1.2 1.2 1.2 .9 .8 1.6 1.2 1 1.3 1.2 1.?
/ .3 1.4 .5 .3 .5  .5 .7 .7
/OU SI=MD1COV.SIG
```

<u>Note</u>

[a]This program is to write a Sigma matrix according to the specified parameter values. The Sigma matrix will be used as the input matrix to produce non-centrality parameters for the intended power analyses.