MANY HELPING HEARTS: AN EVALUATION OF PEER GATEKEEPER TRAINING IN SUICIDE RISK ASSESSMENT

CAROL STUART
Ryerson University, Toronto, Ontario, Canada

JUDITH KELLY WAALEN, PhD
Canadian Memorial Chiropractic College, Toronto, Ontario, Canada

ECHO HAELSTROMM
NEED Crisis and Information Line, Victoria, British Columbia, Canada

Literature reviews on suicide prevention programs have presented conflicting results on the efficacy of school-based prevention programs. Gatekeeper training and peer helping are both recommended as part of a comprehensive school-based prevention program, yet there is no literary evidence of the systematic evaluation of gatekeeper training for peer helpers.

This study evaluated the efficacy of such training with high school peer helpers using a repeated measures design. Significant gains in knowledge about suicide and skills for responding to suicidal peers were evident immediately after training and 3 months later. There was also a significant improvement in positive attitudes toward suicide intervention following training. Although there was no control group, the research offers tentative support for the efficacy of training peer helpers in suicide risk assessment and indicates the importance of additional training for peer helpers.

Research has demonstrated success with training adult populations in the techniques of suicide risk assessment (Haw & Andres, 1998; Ramsey, Tanney, Tierney, & Lang, 1994; Robinson-Ward, 1995; Tierney, 1994). In spite of the fact that Canada and the United States have among the highest rates of youth suicide in the industrialized world (Johnson,
training adolescents as gatekeepers is rare. This study examines whether adolescents can be successfully trained as gatekeepers to identify other adolescents at risk for suicide and refer them to professional resources for crisis intervention.

Several reviews of the evaluation literature on suicide prevention programs have presented conflicting results, either supporting or refuting the efficacy of school-based prevention programs, particularly suicide awareness education programs (Hayden & Lauer, 2000; Lewis & Lewis, 1996). Recommendations regarding school-based suicide prevention programming include both gatekeeper training of adults and peer support programs (Hayden & Lauer, 2000; Kalafat & Elias, 1995; Tierney, 1998). However, there is no evidence in the literature about school-based suicide prevention of gatekeeper training with peer helpers. There are, however, some authors in the literature on peer helping who recommend training peers in specific suicide intervention strategies and in the recognition of suicide risk factors (Herring, 1990; Lewis & Lewis, 1996; Martin, Martin, & Barrett, 1987).

Proponents of peer helping have demonstrated the efficacy of training adolescents in basic helping skills. Adolescents recognized as natural helpers by their peers deliver prevention programs in conflict mediation, tutoring, basic listening and support, substance abuse, assault, and racism (Myrick & Folk, 1991; Tindal & Salmon-White, 1990). In addition to basic helping skills, peer helpers receive training on the above issues and supervision and support with school personnel. Because studies have shown that 25% to 40% of male peers and 40% to 60% of female peers know someone who has attempted suicide, and only about 25% of these confidants actually tell an adult about their suicidal peer (Kalafat & Elias, 1995), a natural extension of peer helping would be training peers in suicide risk assessment. However, evaluations of peer helping programs, both general and specific to suicide prevention, are few and their conclusions are limited because of methodological issues (Lewis & Lewis, 1996).

Kalafat & Elias (1994) demonstrated that, compared with a control group, adolescents who participated in suicide intervention classes made significant gains in relevant knowledge about suicidal peers. They also developed significantly more positive attitudes toward intervening with suicidal peers and a significantly enhanced understanding of the choices for correct action in scenarios involving suicide ideation. These effects, however, were not present for all the items used in the knowledge and
attitude scales, and less than one-third of the items had significant effects. The authors suggest that when designing an instrument for assessing change following training it is important to target the items covered by the curriculum.

Studies on skill development during training have used role-play or simulation; rated participants on generic helping skills such as empathy, warmth, and genuineness; or focused on paper-and-pencil tests of skills development that may or may not test the specific skills related to assessing suicide risk. Problems with this type of research include (a) the lack of inter-rater reliability for assessment of role-play situations and instrument validity for paper-and-pencil tests, (b) a lack of focus on the unique skills of suicide intervention, and (c) logistical issues related to actually taking the time to complete such assessments (Neimeyer & Pfeiffer, 1994).

In summary, reviewing the literature pointed to the lack of studies assessing the use of adolescents as peer gatekeepers for suicide risk assessment. Following on this finding was the lack of tools designed for adolescents to measure the effect of training. Thus, the objectives of this study were two-fold: to assess changes in the knowledge, skills, and attitudes toward suicide among peer helpers as a result of suicide-risk assessment training and to assess the revised Suicide Intervention Response Inventory (SIRI-II) for use with adolescents. Using a repeated measures design, participants were assessed at baseline, immediately after the training, and 3 months after training.

We hypothesized that:

1. Knowledge about suicide would increase after training and be maintained over time.
2. Positive attitudes toward suicide intervention would increase after training and be maintained over time.
3. Skills in suicide assessment and listening skills would become more like those of expert helpers after training and be maintained over time.

**Methods**

This section briefly describes the Peer Gatekeeper Training program offered to eight schools on the south end of Vancouver Island in British Columbia, Canada. Following a description of this suicide prevention program, the research design and procedures are outlined.
Peer Gatekeeper Training in Suicide Risk Assessment

The province of British Columbia, Canada has a well-developed youth suicide prevention effort, which is coordinated through the Suicide Prevention Information and Resource Center (SPIRC). The communities in the large urban-center area where the study occurred had a full continuum of services for youth consistent with the usual recommendations for a comprehensive suicide prevention program (Hayden & Lauer, 2000; Kalafat & Elias, 1995; Tierney, 1998). The local crisis line and information service offered suicide awareness education for students in the 9th–12th grades in three local school districts. The Center received funding from private sources to offer “Many Helping Hearts,” a Peer Gatekeeper Training (PGT) program.1 The Faculty of Education at the local university has offered courses in peer-helping programming and training to teachers, school counselors, and other professionals since the 1980s (Carr, 1981; de Rosenroll, 1990), supporting the local secondary schools to offer peer-helping programs.

The concept of peers as first responders to suicidal ideation in adolescents was already accepted in the community because of the suicide awareness education. The eight schools recruited for the suicide risk assessment training had existing peer helping programs. The training consisted of two half-day sessions, approximately 1 week apart, with a follow-up session 3 months later. The Center’s coordinator of Suicide Prevention delivered training for youth to each group at their home school. The skill-based training sessions included a variety of training techniques covering the following areas: (a) active listening skills, (b) self care and setting limits, (c) crisis theory, (d) signals of suicide, (e) suicide risk assessment, (f) role-play scenarios involving suicidal youth, and (g) community resources.

Sample

The study took place from January to April 2000, recruiting volunteers for the PGT program from among the existing, trained peer helpers. Because this was the first time PGT was offered, a purposeful sampling strategy was used to select a school from each of the three local school districts.

---

1The PGT was funded by the Victoria Real Estate Board and the United Way. Support for the evaluation was provided by the Ryerson University Internal Social Sciences and Humanities Research Council Grants Committee.
districts, a private school, and an alternative school. This ensured representation from students with a variety of backgrounds and personal experiences. Training groups were small (13 students; age range 12–14 years). All participants volunteered to complete the questionnaires.

A total of 65 adolescents, with an average age of 15.6 years (range 13–18 years) participated. The ratio of female to male participants was 3:1 and 80% of the peer helpers were in their first year of a peer-helping program. The majority of the participants (79%) had two or more days of training in basic peer-helping skills, 81% had no prior training in suicide awareness, and 19% had attended a basic suicide awareness education session in the school. Consistent with previously cited research (Kalafat & Elias, 1994), 46% reported contact with a suicidal peer. Thirty-seven participants completed all three administrations of the questionnaire. The low response rate was a function of the way in which the training was scheduled into the school day. Students were excused from their regular classes to attend training and felt pressured to return to their academic schedules. Typical of teenagers, they forgot to return the questionnaires, even with numerous reminders, if they volunteered to complete it outside of class time.

**Instrumentation**

The instrument was developed by selecting items from three instruments used in other studies on suicide awareness and suicide risk assessment training. Consistent with the strategy suggested by Kalafat & Elias (1994), the PGT trainer and her supervisor chose questions from these instruments that were most relevant to the training.

Skills were assessed using items from the SIRI-II (Neimeyer & Bonnelle, 1997). Fifteen of the original 24 questions were selected. In two of the questions, language was modified to reflect the client’s age (e.g., instead of *wife*, the term *girlfriend* was used). Each question on the SIRI-II presents a scenario and two responses, as illustrated in the following example:

Student: I decided to talk to you tonight because I really feel like I might do something to myself . . . I’ve been thinking about suicide.
Helper A: You say you’re suicidal, but what is it that’s really bothering you?
Helper B: Can you tell me more about your suicidal feelings?

Participants rate the responses of both Helper A and B on a 7-point scale ranging from −3 (not appropriate) to +3 (very appropriate). The
scoring procedures outlined by Neimeyer & Bonnelle (1997) require that the participant’s answer be subtracted from the expert score, which is the average of the responses of a panel of experts who completed the SIRI-II during its development. For the example above, the expert score for Helper A is $-2.71$; therefore a participant rating of $-2$ for the response would result in a score of $-0.71$. The total score on the SIRI-II was computed by averaging the absolute value of all responses. The resultant value represented how close the participant responses were to those of an expert. A lower value implied skills that were closer to those of an expert.

The SIRI-II has been rigorously assessed for validity and reliability by the authors. Neimeyer and Bonnelle (1997) reported test—retest reliability co-efficient of $.92$ ($p < .001$) and strong internal consistency with co-efficient alphas of $.90$ (pre-test) and $.93$ (post-test). For this study, the internal consistency was similar to the original with coefficient alphas of $.88$ (pre-test) and $.86$ (post-test).

An additional modification to the SIRI-II asked participants to provide open, unstructured responses to questions 1, 3, 9, 13, and 21, in addition to scoring the responses. The intent was to capture the natural responses of adolescents for scenarios with varying degrees of suicide ideation and symptoms. These responses were analyzed qualitatively.

Ten items from the Suicide Intervention Questionnaire (SIQ), previously used in other suicide intervention training programs, assessed changes in attitudes toward suicide intervention (Haw & Andres, 1998; Tierney, 1994). As illustrated in the following example each statement is rated by the participant using a 5-point Likert scale ranging from 5 (strongly agree) to 1 (strongly disagree):

\[
\text{I have the right to intervene if someone is at risk of suicide. } 1 \quad 2 \quad 3 \quad 4 \quad 5
\]

Three of the questions (2, 8, 26) have a reverse valance relative to the desirable attitudes. A higher mean score represents more favorable attitudes toward suicide intervention. The original 20-item questionnaire for the SIQ has a split-half reliability coefficient of .79 (Pearson) and a test-retest Pearson correlation coefficient of .82. The modified SIQ had a moderate test—retest reliability ($r = .67, p = .01$) and a moderate split-half reliability of $r = .67$ (Guttman).

The knowledge questions consisted of eight true-and-false questions about suicide, one question asking the participant to list three signals of
suicide, and one multiple choice question to identify his/her main responsibility in supporting a suicidal teen. For example:

Dramatic changes in behavior may be a signal for a suicide attempt. T F

These questions were used in the evaluation of a suicide awareness education program that was delivered annually by the Center to students in the 9th to 12th grades (Stuart & Kline, 1996). The responses were totalled for the knowledge score, with a maximum of 12. The instrument also included a set of questions about satisfaction with training and background information on demographic characteristics and peer-helping experiences.

**Data Collection**

The questionnaire was administered before the training, immediately after the completion of training, and 3 months following training. Informed consent was obtained prior to the pre-test, following the guidelines of the Ryerson University Research Ethics Board. Confidentiality was maintained using an identification number and face sheet with the student’s name, which was removed once the next questionnaire in the sequence was distributed. The pre-test was distributed by the school counselor one week prior to the first training and collected by the trainer at the first session. The trainer administered the post-training questionnaire after the second training session. The school counselors distributed the follow-up questionnaire and it was collected at the beginning of the session by the trainer. In an effort to maximize the response rate, questionnaires for the trainees not at the follow-up were left with the school counselor who distributed, collected, and returned the forms.

**Data Analysis**

Results were analyzed separately for each of the questionnaire’s three subscales, using SPSS (v. 9.0). To examine change over time, repeated measure analyses of variance (ANOVAs) were calculated for skill, attitude, and knowledge scores. An alpha level of .01 was used for all statistical tests. If the main effect was significant, then pair-wise comparisons were examined to determine the significance of the training effect and whether change occurred in the predicted direction and was maintained over time.
The unstructured responses on the SIRI-II were entered into text files and imported into ATLAS-Ti (Muhr, 1998). Carol Stuart coded the responses of all 65 participants according to the type of response (e.g., open question, supportive, reflection of feeling, etc.) and the focus of the response (e.g., defending the respondent, suicide ideation, advice, solution). The responses and the codes were reviewed, looking for commonalities and differences within each question and across each administration of the questionnaire.

**Results**

Because the sample was drawn from five schools, each with potential differences in the characteristics of their peer-helping programs, a test for interaction effects based on these differences was conducted. There were no statistically significant differences between the participants due to the level of previous peer helper training or level of support provided to peer helpers, therefore the five schools were treated as a single group for the purposes of testing the hypotheses.

As indicated in Table 1, the repeated measures ANOVA was significant for all three dependent variables (knowledge, attitudes, skills), indicating that at least one mean difference was significant for each variable. The pair-wise comparisons for the skill measure indicated that the

<table>
<thead>
<tr>
<th>TABLE 1</th>
<th>Repeated Measures Analysis of Variance (ANOVA) for Knowledge, Skill and Attitude Measures</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre-test</td>
</tr>
<tr>
<td>---------</td>
<td>----------</td>
</tr>
<tr>
<td></td>
<td>M</td>
</tr>
<tr>
<td>Knowledge</td>
<td>9.86</td>
</tr>
<tr>
<td>Attitudes (SIQ)</td>
<td>3.77</td>
</tr>
<tr>
<td>Skills (SIRI-II)*</td>
<td>1.65</td>
</tr>
</tbody>
</table>

*The mean represents the participant’s deviation from the expert score; therefore a smaller number is a better score.

Note: SIQ = Suicide Intervention Questionnaire; SIRI-II = Suicide Intervention Response Inventory.
pre-test scores differed significantly from the post-test (mean difference = .475, \(p = .0001\)) and the follow-up test (mean difference = .471, \(p = .0001\)). Both sets of post-training scores were significantly different from the pre-test, supporting the hypothesis that the skill level is more like that of the experts after training than before. The difference between the post-test and the follow-up test was not significant, supporting that hypothesis that the change in skills was maintained over time. The pairwise comparisons for the attitude measure indicated that the pre-test scores differed significantly from the post-test (mean difference = -.574, \(p = .0001\)) and the follow-up test (mean difference = -.348, \(p = .0001\)) in the direction of improvement in attitudes toward suicide intervention after training. The difference between the post-test and the follow-up test was also significant (mean difference = .216, \(p = .001\)) in the direction of a loss in favorable attitudes. The results support the hypothesis that positive attitudes toward suicide intervention increased after training. The hypothesis that change over time would be maintained was not supported, however there was still a significant difference in attitudes between the pre-test and the follow-up test. The pairwise comparisons for the knowledge measure indicated that the pre-test scores differed significantly from the post-test (mean difference = -.757, \(p = .002\)) and the follow-up test (mean difference = -.811, \(p = .003\)). The difference between the post-test and the follow-up test was not significant. The results support the hypothesis that knowledge increased after training and was maintained 3 months later.

The next section presents some observations from the analysis of the open-ended responses to enhance the understanding of the scored responses to the SIRI-II presented above. Two of the questions requiring open-ended responses had direct references to suicide in the scenario. On both of these questions there was a post-training increase in the participant’s ability to formulate questions focused on the suicidal intent. Pre-training, 12% of the responses for Question 1 and 6% of the responses for Question 21 (original numbering from the SIRI-II) inquired about suicide ideation. As this student response indicates, these pre-training responses typically searched for the source of the suicidal feeling: “You are such a great person and friend. What would make you think about suicide? Why are you feeling this way?” Students were more capable of inquiring about suicide ideation after they had completed the PGT program. Immediately post-training, 46% of responses to Question 1 and 32% of responses to Question 21 inquired about suicide.
A typical post-training response also included a paraphrase of feelings similar to the following response: “It sounds like you are feeling nobody cares about you. Have you ever thought about suicide?”

The other three questions asking for open-ended responses (3, 9, 13 from the original SIRI-II) contained more subtle signs such as anger, boredom, and irrational thinking. Pre-training, student responses to these questions tended to provide advice or solutions, similar to the following responses: “You can talk freely with me, everything we discuss stays between us, you can tell me about your thoughts” (Question 3); “Life is hard no matter if you are rich or poor, sometimes you just need someone to talk to, that’s why I’m here” (Question 13). Basic peer-helping training focuses on reflecting feelings and discourages early solutions and advice. Having already completed basic training, these students should have been able to focus on reflective responses involving feeling. Very few of them did this. Reflection of feelings was evident in 15% to 23% of the pre-training responses to these questions. In post-training the rate of reflection of feeling went up to 41% to 50%. The other skills covered in peer-helping training such as open-and-closed questions, reflection of content, and exploring the issue saw little change on the open-ended responses throughout the three administrations.

**Discussion**

Peer helping has been identified as one component of a well-developed school-based suicide prevention strategy. There has been little indication in the literature of either the need for, or benefit of, training peers as gatekeepers. This study offers strong preliminary support for such training. Significant differences in skills, attitudes toward suicide intervention, and knowledge occurred after training and were maintained in all areas over a 3-month period of time.

Training peer helpers in general helping skills and expecting that they will be effective in a school-based suicide prevention program is naïve. They need skill-specific training for suicide risk assessment. As noted earlier, only 25% of peer confidants will tell an adult about a friend’s suicide ideation (Kalafat & Elias, 1995). The results of the SIRI-II support this assertion. Even though they were trained peer helpers, they did not have suicide intervention skills. These skills were developed and maintained with specific training.
The modified SIRI-II seems to be a useful tool for measuring skill development with adolescents as it stands. After training, participants became more like the experts in their choice of response. The open-ended responses to the SIRI-II illustrate the dramatic shift in participants’ ability to ask specific questions about suicide ideation and indicate areas where training may need to be improved. The more subtle signs of potential suicide ideation and depression tended to be missed by participants, even after training. It may be advantageous to discuss and give examples of these types of clues during training. In particular, the angry and accusatory statements of a peer that could indicate underlying depression are difficult for adolescents to understand and respond to. Without specific training in what such statements represent, peer helpers could distance themselves from such students, effectively increasing the risk for suicide as the student becomes more isolated.

Although the results are encouraging, the reader is cautioned to keep in mind the limitations of this study. It is possible that the results could be explained by other events that occurred over the time period of this study. Because there was no control group of peer helpers who did not receive suicide risk assessment training, we can only tentatively attribute the results to the training. Although there were 65 participants in total, only 37 completed all three administrations of the questionnaire and it is possible that these 37 had different characteristics from those that did not complete all three test administrations. Future research of this nature might consider a control group and include a larger sample in order to overcome these potential limitations.

It is also important to assess the transfer to training in suicide risk assessment to the natural environment by tracking the referrals related to suicide ideation that are a direct result of peer helper involvement. Teacher/counselor supervisors can be used as co-researchers with an expectation that, where possible, peer helpers will accompany a peer they are concerned about to an initial meeting with the teacher/counselor. The teacher/counselor would then be in a position to observe the peer helper’s skills and, after completing their own risk assessment, compare the assessments.

This research offers tentative support for the efficacy of training peer helpers in suicide risk assessment and indicates the potential they have to offer in a comprehensive suicide prevention program in the schools. This component of school-based suicide prevention needs to go beyond the strategy of a basic peer-helping program. For peer helpers to be an
effective component of a suicide prevention program they not only need basic training in empathy and active listening, but they also need training in suicide risk assessment and encouragement to seek out the angry and isolated peers in their school and develop relationships with them even though those relationships might be resisted or difficult.

References


