Preventing Suicides in the Toronto Subway System: A program Evaluation

By

Rahel Eynan

A thesis submitted in conformity with the requirements for the degree of

Doctor of Philosophy

Institute of Medical Science

University of Toronto

© Copyright by Rahel Eynan (2011)
Preventing Suicides in the Toronto Subway System: A Program Evaluation

Rahel Eynan

Institute of Medical Science, University of Toronto

ABSTRACT

Despite the wealth of information on suicide prevention issues and the widespread implementation of suicide prevention strategies, program evaluation efforts have been limited. Lack of sound program evaluation remains one of the most significant barriers to identification and implementation of effective intervention and prevention strategies. The purpose of this study was two-fold: to conduct a summative evaluation of the gatekeeper suicide prevention program implemented at the Toronto Transit Commission, and to concomitantly, appraise the efficacy and effectiveness of the Kirkpatrick evaluation model as an analytical framework to guide suicide prevention program evaluations. The study used a two-phase, sequential mixed-method approach of converging quantitative and qualitative methodologies. The quantitative study employed a repeated measures design and examined the immediate and long-term effects of the gatekeeper program on attitudes, knowledge, intervention abilities. The qualitative study consisted of semi-structured interviews and explored participants’ effective and utility reactions to the gatekeeper training program. The results of this study indicated safeTALK and suicideAWARE training programs increased participants’ knowledge of suicide and suicidal behaviour, enhanced positive attitudes toward the suicidal individual, suicide intervention, and improved intervention skills. The empirical findings from this study support the premise that the Kirkpatrick evaluation model could be adapted for use in gatekeeper program evaluations. The model provides a highly relevant, well-rounded, rigorous approach to suicide prevention program evaluations.
ACKNOWLEDGEMENTS

The completion of this dissertation would not have been possible without the support of many people to whom I remain deeply indebted. A special thanks to my academic supervisor, Professor Paul Links. I appreciate his consistent support, generosity, diligent commitment to this research, his warm and humble nature, and the wealth of knowledge he shared throughout this process. It was a privilege working with him. I also would like to thank my Program Advisory Committee members: Dr. Carol Strike, Dr. Janet Durbin, Dr. Ronald Heslegrave, and John O’Grady who contributed their wisdom and time throughout the completion of this research. They have offered support and guidance while challenging me to achieve, both in the process of this dissertation and in the years prior to its completion. To my dissertation reviewers: Dr. Flora Matheson, Dr. Paul Kurdyak, Dr. Mary Seeman, and Dr. Barry Hall for their insight and constructive criticism.

Those to whom I owe my greatest debt have been promised anonymity. They are the Toronto Transit Commission employees who have volunteered their time and effort towards this study. A special thanks to those employees who have agreed to be interviewed and have candidly shared their experiences. They all have my heartfelt gratitude.

A special thanks to my colleagues Jesmin Antony, Yvonne Bergmans, Debbie Ezard, and Dr. Anne Rhodes at the Suicide Studies Research Unit at St. Michael’s Hospital for their continued support, thoughtfulness, friendship, and sense of humour. Many thanks also to Judy James for her attention and assistance which far exceeded my expectations.

My deepest thanks to my family and friends for their continued support, love, and encouragement. Most importantly, I need to richly and deeply thank my children, Ravit and Alexandre, for believing in me implicitly, and for supporting me relentlessly through this journey. Their support has allowed me to follow my dream. They have anchored me through the times when I wavered and have celebrated in the victories along the way. I would not have been able to do this without their support, patience, and love.

I dedicate this work to Libi Kayla and Shai Eden, my grand-daughters, whom I love all the way to the moon and back, ten times.
# TABLE OF CONTENTS

Abstract ................................................................................................................................. ii
Acknowledgements ............................................................................................................... iii
Table of Contents ............................................................................................................... iv
List of Tables ..................................................................................................................... v
List of Figures .................................................................................................................... vii
List of Appendices ............................................................................................................ viii
Introduction ...................................................................................................................... 1
Subway Suicides .............................................................................................................. 14
Program Evaluation and Theoretical Approaches ............................................................ 27
Methods ............................................................................................................................. 52
Quantitative Results ......................................................................................................... 90
Qualitative Results ........................................................................................................... 128
Discussion and Conclusions ............................................................................................ 157
References ......................................................................................................................... 184
LIST OF TABLES

Table 1. Summary of The Kirkpatrick Model (TKM) ......................................................... 56
Table 2. Summary of Participation in the Program Evaluation .............................................. 60
Table 3. Characteristics of 305 Workshop Participants ...................................................... 61
Table 4. Sample Characteristics of safe TALK Workshop .................................................. 62
Table 5: Sample Characteristic of suicideAWARE Worship participants .......................... 63
Table 6. List of Outcome Measures Used in the Evaluation ............................................... 65
Table 7. Summary of Targeted Occupational Groups, Workshops and Evaluation Measures .................................................................................................................. 65
Table 8. Phase two Sample characteristics ........................................................................ 81
Table 9. Sample Characteristics of the 307 Participants from Whom Demographic Data Was Gathered ........................................................................................................ 93
Table 10. Workshop Participants’ Experience with Suicidal Behaviour While with the TTC ................................................................. 95
Table 11. Workshop’s Participants’ Experience with MHA Apprehension While with the TTC ................................................................. 97
Table 12. Comparison of Proportions of Correct Responses on the Measures ................. 100
Table 13. Immediate and Long-term Effects of the Training on Estimated Marginal Mean Scores .................................................................................................................. 101
Table 14. Immediate and Long-term Effects of Training on Raw Mean Scores of Outcome Measures ........................................................................................................ 102
Table 15. Comparison of the Proportions of SPQ Correct Responses of the Occupational Groups ........................................................................................................ 103
Table 16. Immediate and Long-term Effects of Training on the SPQ Raw
Mean Scores of the Occupational .................................................................105

Table 17. SPQ Model Based on Differences .....................................................109

Table 18. Comparison of Proportions of IKT-R Correct Responses of the
Occupational Groups .....................................................................................111

Table 19. Comparison by Occupational Group of the Effects of Training on
IKT-R Raw Mean Scores ..............................................................................113

Table 20. IKT-R Model Based on Differences ................................................114

Table 21. Comparison of the Proportions of SOQ-R Correct Responses of the
Occupational Groups .....................................................................................118

Table 22. The Effects of the Training on the SOQ-R Raw Mean Scores of the
Various Occupational Groups .......................................................................119

Table 23. SOQ-R Model Based on Differences ................................................121

Table 24. Comparison of Proportions of SIRI-R Correct Responses of the
Occupational Groups .....................................................................................125

Table 25. SIRI-R Raw Means by Occupational Group ....................................126

Table 26. SIRI-R Model Based on Differences ................................................127
LIST OF FIGURES

Figure 1: The Kirkpatrick Model Pyramid .......................................................13

Figure 2: The interaction between occupational group affiliation and the immediate and long-term effects of the training on SPQ mean scores……107

Figure 3: The interaction between occupational group affiliation and immediate and long-term training effects on the SOQ-R mean scores………………120
LIST OF APPENDICES

Appendix A: The Collaborative Project ................................................................. 203
Appendix B: The development of a comprehensive warning sign list ...................... 206
Appendix C: Logic Models for the Program Evaluation ............................................. 209
Appendix D: suicideAWARE pre-training Questionnaire ........................................ 212
Appendix E: safeTALK pre-training Questionnaire .................................................. 218
Appendix F: suicideAWARE post-training Questionnaire ....................................... 228
Appendix G: safeTALK post-training Questionnaire .............................................. 232
Appendix H: suicideAWARE 3-months follow-up-training Questionnaire ............... 240
Appendix I: safeTALK 3-months follow-up Questionnaire ..................................... 245
Appendix J: Qualitative interview guides ............................................................... 254
Appendix K: Effect Size (ES) formula .................................................................. 278
Appendix L: Consent forms .................................................................................... 280
Appendix M: List of Community Resources .......................................................... 285
Appendix N: Linked and Unlinked Tables ............................................................... 287
CHAPTER ONE
INTRODUCTION

Suicide is a significant global public health problem, and ranks among the top ten causes of death in Western countries (Canadian Association for Suicide Prevention (CASP), 2004; Diekstra, 1989; Sakinofsky, 1998; U.S. Public Health Service, 1999; World Health Organization, 2000). Beyond the surveillance data that likely represents an underestimate of the actual number of individuals who harm or kill themselves (Leenaars, 1995) lies a broad spectrum of economic, personal and risk factors that are indicative of the imperative need for suicide prevention programs (Bongar & Harmatz, 1991; CASP, 2004; Tanney, 1989; Tierney, 1994, 1988; U.S. Public Health Service, 1999; WHO, 2000).

Over the past two decades, the first two generations of suicide prevention efforts have generated voluminous information on risk and protective factors, empirically-based methods for the prevention of suicidal behavior, and improved research methods (Berman & Jobes, 1995; Public Health Service, 2001). The impetus for suicide prevention training programs emanated from an emergence of research that identified skills and strategies that may facilitate preventative intervention (Eggert, Thompson, Herting, & Nicholas, 1995; Thompson, Eggert, Randell, & Pike, 2001). The proliferation of curriculum-based suicide prevention programs in schools (Garland, Shaffer, & Whittle, 1989) concomitant with an increased attention and concerns voiced over format, goals, theoretic orientation, and safety issues (Hazell & King, 1996) led to improved methods and prevention program designs (Breton et al., 2002; Kalafat & Ryerson, 1999) and to a more precisely defined prevention framework that placed prevention programs on a continuum of universal, selective and indicated interventions (Gordon, 1987; Institute of Medicine, 1994).

This framework focuses attention on defined populations—from an entire population, to specific at-risk groups, to specific high-risk individuals—i.e., three population groups for whom the designed interventions are deemed optimal for achieving the unique goals of each prevention type.

Universal strategies address an entire population (community or school). These prevention programs are designed to influence everyone, reducing suicide risk through removing
barriers to care, enhancing knowledge of what to do and say to help suicidal individuals, increasing access to help, and strengthening protective processes like social supports and coping skills. Universal strategies include programs such as public education campaigns, school-based “suicide awareness” programs, means restrictions, education programs for the media on reporting practices related to suicide, and school-based crisis response plans and teams. (Gordon, 1987; Institute of Medicine, 1994).

Selective strategies are designed to focus on subsets of the total population: the at-risk groups that have a greater probability of becoming suicidal e.g., persons with mental illness and/or substance abuse,. Selective prevention strategies include screening programs, gatekeeper training for “frontline” adult caregivers and peer “natural helpers,” support and skill-building groups for at-risk groups in the population (e.g., Aboriginal communities, persons in custody, Gay, Lesbian, Bisexual, and Transgender(GLBT) youth) enhanced accessible crisis services and referral sources. (Gordon, 1987; Institute of Medicine, 1994).

Indicated strategies address specific high-risk individuals within the populations—those evidencing early signs of suicide potential. Programs are designed and delivered in groups or individually to reduce risk factors and increase protective factors. Indicated strategies include such programs as skill-building support groups in high schools and colleges, parent support training programs, case management for individual high-risk youth at school, and referral sources for crisis intervention and treatment. (Gordon, 1987; Institute of Medicine, 1994).

**Gatekeeper Programs**

Community gatekeeper training programs, which are considered selective high-risk interventions, have been regarded as a central element of suicide prevention strategies for the past two decades (Mann et al., 2005; Potter, Powell, & Kachur, 1995) and are included in many national suicide prevention strategies (CASP 2004; U.S. Public Health Service, 1999).

Over three decades have passed since Dr. John Snyder (1971) published the first report on gatekeeper training in the Bulletin of Suicidology in which he discussed the management of crisis within a community context and defined the term “gatekeeper.” Gatekeepers are individuals who have primary contact with persons who are at risk of suicide and are able to identify them by recognizing suicidal risk warning signs. Gatekeepers are divided into two categories: “designated gatekeeper” and “emergent gatekeeper.” The latter refers to any
individual who has potential contact with the person at risk and who may not have been formally trained to intervene with someone at risk for suicide i.e.: family members, friends, clergy, teachers, police, coaches, recreational staff, and counsellors, while the former is more clearly associated with a trained professional in a helping role i.e: Social Worker, pharmacist, GP, therapist (Potter et al., 1995).

Providing accessible help is a key feature of any effective community-based suicide intervention strategy. Most communities promote services which provide informal support and/or professional help for individuals who experience suicide-related crisis; however, many potential beneficiaries of these services either are unaware of their existence or are reluctant to access help on their own initiative. Gatekeeper training addresses this challenge by widening the circle of people able to recognize and reach out to individuals at risk. The gatekeeper programs provide educational programs which sensitize and train community members to become effective first responders to individuals who are potentially at risk, linking them with further help.

Thus, gatekeeper suicide prevention programs aim at building competence and confidence to:

- Recognize risk factors associated with suicide;
- Identify at-risk individuals;
- Communicate with at-risk individuals and offer support; and refer the identified high at-risk individual to the appropriate community resources.

Gatekeeper programs are designed to achieve some or all of the following goals:

- To dispel myths and increase knowledge. The programs present facts, statistics, and myths regarding suicide to help participants understand why some people become suicidal and what they, as gatekeepers, can do to prevent suicide.
- To promote identification of potentially suicidal individuals. General suicide education programs provide descriptions of warning signs of suicide and encourage participants to seek help for individuals who are contemplating suicide.
- To encourage participants to seek help for distressed at-risk persons. The programs describe methods of seeking and accessing help.
• To promote the development of suicide intervention skills. Programs also promote the development of listening and interpersonal skills; for example, to help students improve their relationships with peers, parents, and others.

• To provide participants with information about mental health resources. Programs provide participants with information about how various mental health resources operate and how to contact them for referrals.

If successful, gatekeeper programs would presumably result in an increase in identification of at-risk of suicide patrons and their referral to appropriate mental health resources.

The Need for Suicide Prevention Program Evaluations
Despite the wealth of information on prevention issues and the widespread implementation of prevention strategies, there is precious little evidence to support the efficacy of suicide prevention programs in reducing suicide. The effectiveness of suicide prevention programs has not been demonstrated. The need for program evaluation in suicide prevention is well documented (Breton et al., 2002; CASP, 2004; Center for Disease Control [CDC], 1994; Goldney, 1998a, 1998b; Tierney, 1994; US Department of Health and Human Services, 2001). Evaluation efforts have been very limited. No comprehensive, standardized models of evaluation have been developed, whereby training and the effects of various forms of training in suicide prevention can be measured and/or compared.

The National Task Force on Suicide (1987, p. 44) found that evaluation studies were few and those that have been conducted proved to be methodologically inadequate, or based on systems of data collection which could not be compared across services due to differences in behavioural definitions, and selection and measurement of independent and outcome variables. Lack of sound evaluation remains one of the most significant barriers to identification and implementation of effective intervention strategies. Evaluation studies have the potential to produce information on program efficacy and effectiveness and to provide information that will improve program delivery. Incorporating evaluation efforts into suicide prevention programs is imperative. Planning, process, and outcome evaluation are important components of any public health effort to reduce suicide. (CDC, 1994; Goldney, 1998b). Stein and Lambert (1984) subsequent to reviewing the development of program evaluation in suicidology, concluded that suicide prevention program evaluations need to; 1. Cross-validate indices of effectiveness; 2.
Use more than one index of effectiveness in any evaluation; 3. Assess larger, more broadly-based samples.

While the rationale for gatekeeper programs, the need for gatekeeper programs, and the development of gatekeeper programs can be extensively documented, evaluation results cannot. Tierney (1994), in his review of training evaluation efforts noted that program evaluations have largely consisted of trainee and trainer ratings of satisfaction that have been too program specific and methodologically deficient. More recently, in a review of evaluative research on suicide prevention across Canada, Breton and colleagues (2002) found fifteen studies, most of which assessed the effectiveness of school-based programs. While, most evaluations found benefits from the programs, few studies had been published in the scientific literature. The report concluded that there is much need for further evaluative research and that it should focus not only on effectiveness, but also examine the appropriateness of interventions, and their impact both on the specific factors that have been targeted and on the wider health care and social systems. The report also noted that the tendency to focus on one sector (e.g.; school) did not reflect the reality of youths’ experience which cuts across different professional domains and thus, demanded a more integrated and ecological approach.

A recent systematic review by Mann et al. (2005) examined the efficacy of suicide prevention strategies worldwide. The most promising interventions are physicians’ education, means restrictions, and gatekeepers’ education. To date, the systemic evaluation of gatekeeper suicide prevention programs’ impact on suicide rates has been limited to multilevel programs in institutional settings such as military forces. Subsequently, Mann et al. (2005) recommended that evaluations of gatekeeper programs with intermediate outcome measures such as referral rates and psychiatric treatment rates ought to be conducted.

The Suicide Prevention Initiatives and the Implementation of the Gatekeeper Program

Suicides and attempted suicides are an unfortunate reality of modern urban rapid transit systems and have been documented in cities worldwide (Guggenheim & Weisman, 1972; Ladwig & Baumert, 2004; Mishara, 1999; O’Donnell & Farmer, 1992; O’Grady & Griesi, 2002; Sohier & Sutton, 2002). The structural environment of the subway system provides a distinct milieu within which incidents of violence and injuries can occur. One type of violent event
uniquely associated with subways is attempted suicide or death by suicide via jumping into the path of an oncoming train, lying on the tracks or throwing oneself onto the electrified track.

In addition to the obvious loss of human life, suicides on the subway system have traumatic consequences for both transit employees and passengers; it places the public at risk due to train delays and overcrowding on narrow platforms, and has corporate and societal economic ramifications (Ladwig & Baumert, 2004).

The Toronto Transit Commission (TTC) is a public transport authority that operates an extensive network of buses, subway trains, streetcars, light rail vehicles and para-transit buses in Toronto, Ontario, Canada. The TTC operates the third most heavily used urban mass transit system in North America. In 2007, the TTC carried 1.5 million passengers daily, and of those 945,000 were subway passengers. The TTC operates three subway lines and one elevated rapid transit with a total of 69 stations as well as 149 connecting surface routes (buses and streetcars) of which 148 routes make 243 connections with a subway or rapid transit station during weekdays rush hours. The average daily ridership exceeds 2.49 million passengers: 197,000 through buses, 328,700 by streetcar, 35,300 by intermediate rail, and 901,400 by subway. The TTC also provides door-to-door services for persons with physical disabilities known as Wheel-Trans (American Public Transit Association, 2007)

The suicide prevention initiatives at the TTC consisted of several distinct phases and spanned over the past four decades:

Phase 1: Decision by TTC management to enhance their existing suicide prevention strategy which was initiated in the early 1970s. The existing strategy used a three prong approach:

1) Containing, to the extent possible, the flow of information on suicide incidences, and collaboration with the police, coroner’s office, and the local media, community and suicide prevention agencies.
2) Developing well-defined procedures and training of Special Constables in mental health interventions.
3) Instituting engineering approaches to reduce risk i.e., emergency power- cut switches on platforms, safe refuge areas.

Phase 2: Planning and organization of the collaborative project with the Arthur Sommer Rotenberg Chair in Suicide Studies. A detailed description of the Collaborative project can be found in Appendix A.
Phase 3: The development of a comprehensive list of warning signs; the development of the warning signs list is described in Appendix B.

Phase 4: This suicide prevention initiative at the TTC commenced in 2005 with the implementation of the gatekeeper suicide prevention program. The gatekeeper training program was offered to TTC Special Constables, Mobile, Route and Subway Supervisors, and Train Operators.

Phase 5: The present gatekeeper program evaluation constitutes the last phase of the suicide prevention initiative.

More recently, TTC has expanded the gatekeeper program to other employee groups such as ticket collectors.

**The Gatekeeper Suicide Prevention Program Implemented at the TTC**

Overall, the TTC management was committed to the sensitization of all TTC personnel to suicide prevention. The rationale for the implementation of a gatekeeper suicide prevention program at the TTC is illustrated in the Logic Model (Appendix C). The basic premise was that the more TTC employees know about suicide warning signs and resources for help, the more likely they are to identify at-risk patrons and the more likely they will be to get involved and intervene.

The Gatekeeper Training Program had three essential elements:

- The identification of staff and their appropriate roles in gatekeeper training and the establishment of warning signs to screen for distressed persons at risk for suicide.
- The hiring of professional trainers with experience and expertise in gatekeeper training programs. These professionals will implement the three levels of training.
- The development of an evaluation and feedback process to inform the TTC and other stakeholders and other transit authorities about the value of the program.

Through a Request for Proposal (RFP) process Trillium Health Centre won the bid to deliver the educational intervention. LivingWorks Education of Alberta in partnership with Trillium Health Centre developed the suicideAWARE and safeTALK. The training workshop consisted of various training approaches that were designed to target different occupational groups of TTC personnel and were specific to their level of contact and involvement with distressed individuals at-risk of suicide. Both workshops were designed to assist TTC employees in the development
of attitudes, knowledge, and skills to recognize individuals at risk of suicide; however, *safeTALK* also provided workshop participants with assessment and intervention competencies. The workshops were intended to be appropriate for Train Operators, TTC Special Constables, and Transportation Supervisors. A detailed description of the workshops is provided in Chapter Four.

The final goal of the collaborative agreement was to develop an approach to an evaluation project to determine the success of the gatekeeper program from the employees’ perspective. It was determined that the program evaluation should have both specific and broad implications for the advancement of suicide prevention. Specifically, the program will provide feedback to the TTC about the effectiveness of the Gatekeeper Training Program in sensitizing staff to the needs of individuals at risk for suicide and to staff roles in preventing suicides.

The suicide prevention program was delivered in the following manner:

1) Suicide awareness pamphlet: the pamphlet described the magnitude of the problem, dispelled suicide myths, listed warning signs, and listed resources. The suicide awareness pamphlets were attached to salary stubs and distributed to all TTC personnel.

2) Trillium Health Centre in collaboration with LivingWorks Education developed two gatekeeper suicide prevention programs tailored to address the training needs of various occupational groups within the TTC. The training programs were of different formats and delivery methods:

- The three-hour long *suicideAWARE* workshop was delivered to train operators as a component of their re-certification training;
- A one-day *safeTALK* workshop was delivered to: TTC Special Constables, Mobile, Route and Subway Supervisors, Chief Supervisors, training Instructors and safety personnel.

A more detailed description of the workshops is provided in Chapter Four.

**The Gatekeeper Suicide Prevention Program Evaluation**

The outcomes of the gatekeeper suicide prevention training program implemented at the Toronto Transit Commission (TTC) are the primary interest of this program evaluation aimed at determining whether the gatekeeper program objectives were achieved within the institutional context for which they were intended.
The field of suicidology has little to offer regarding the development and application of useful models of gatekeeper program evaluation. Consequently, the second purpose of the study was to examine the effectiveness and relevance of The Kirkpatrick Model (Kirkpatrick, 1959a; 1959b; 1960a; 1960b; 1996; 1998) as a guiding framework for the evaluation of gatekeeper program efficacy.

Rationale for the Summative Evaluation of the TTC Gatekeeper Program

Program evaluation is the systematic collection of information about the activities, characteristics, and outcomes of programs to allow informed judgments about program improvement, program effectiveness, and decisions about future programming. The three primary purposes of evaluation are program planning, program development, and program accountability. The purpose of an evaluation and not the timing determines the type of an evaluation to be conducted (Chambers, 1994). It is the use which is made of the information gathered that determines whether an evaluation is summative or formative.

Formative evaluations aim to describe how the program is actually functioning and are conducted during the development or ongoing implementation phase of a program with the intent to improve the program. Summative evaluations aim at delineating clearly the benefits generated by the program at the end of the program activities (summation). The focus in summative evaluations is on the outcome. Outcomes are usually in terms of enhanced learning (knowledge, perceptions/attitudes or skills) or conditions, e.g., confidence, competence, self-reliance (Patton, 1990a;1990b).

The intent of the evaluation of the gatekeeper program implemented at the TTC was not to influence program innovation, process, or format; but, rather, to evaluate the gatekeeper program against the success criteria based on the following broad objectives traditionally used to evaluate such programs:

a) The degree to which the program implemented sensitizes gatekeepers to their roles in preventing suicide,

b) The extent to which the program results in the appropriate identification and referral of distressed individuals,

c) The impact of the program on suicide incidents.

Consequently, it was the purpose of the TTC evaluation that established the necessary conditions
for defining the evaluation.

The summative evaluation was guided by the following hypotheses:

1. Factual knowledge about suicide and risk factors for suicide would increase after the training and would be maintained over time.
2. Positive attitudes towards suicide intervention would increase after training and would be maintained over time.
3. Suicide assessment and intervention skills would improve and be maintained over time.

The Program Evaluation Questions

A series of questions were raised for the program evaluation. Since the objectives were related to three main areas of attitudes, knowledge, and intervention skills, the evaluation questions focused specifically on these three aspects of the training. Questions and sub-questions which the evaluation attempted to answer can be organized under two major sub-headings:

1. Are the suicide prevention program objectives achieved?
   a) Do participants’ attitudes towards suicide prevention change as a result of the workshop?
   b) Do participants acquire knowledge?
   c) Are participants’ interventions skills enhanced by the training, and are participants able to apply the intervention skills?
   d) Are the enhanced attitudes, acquired knowledge, and intervention skills retained over time?

   It was considered appropriate in the scope of this evaluation to focus mainly on the immediate and long term effects (3 months follow up) of the workshop.

   The program evaluation used a two-phase, sequential mixed-method approach (Creswell, 2003) of converging quantitative and qualitative methodologies: a) a quantitative study of the effect of the gatekeeper program on attitudes, knowledge, intervention abilities, and b) a qualitative study of TTC Special Constables and Subway Supervisors’ satisfaction with the program and its impact on their competence, confidence and willingness to engage and intervene with individuals at-risk.

   There is a growing consensus among evaluation experts that both qualitative and quantitative methods have a place in the performance of effective evaluations. The range of
possible benefits that carefully crafted mixed method designs can yield has been conceptualized by a number of evaluators (Greene, Caracelli, & Graham, 1989; Miles & Huberman, 1994; Patton, 1990a) who purported that both formative and summative evaluations are enriched by a mixed method approach.

**Theoretical Framework: The Kirkpatrick Model (TKM)**

The application of The Kirkpatrick Model (TKM) as a theoretical framework guiding the evaluation allowed the testing of its effectiveness as an evaluative tool of gatekeeper program. The evaluation literature has long recognized the need for comprehensive, systematic evaluation of the effectiveness of training with respect to increasing both training–related knowledge and performance. By far, the most widely used model for evaluating training programs’ outcomes was proposed in 1959 by Donald L. Kirkpatrick. A recent study indicated that 94 percent of the companies surveyed use some form of the Kirkpatrick Model to evaluate their training and development programs (Bassi, Benson, & Cheney, 1996). While Kirkpatrick’s four-level model continues to be the most influential and prevalent framework for categorizing training criteria among researchers (Kirkpatrick, 1996) it is rarely fully implemented. Most program evaluations are usually confined to the “happiness sheets” i.e., trainee reactions (Kirkpatrick, 1998, p. 25).

The simple taxonomy of training criteria addresses the need to understand training evaluation simply yet systematically (Shelton & Alliger, 1993). As indicated in Figure 1, The Kirkpatrick Model (TKM) focuses on the outputs of training by establishing four levels that represent a sequence of measures to evaluate the outcome of any training program (Kirkpatrick, 1959a; 1959b; 1960a; 1960b; 1996; 1998). Each successive level increases in complexity and represents a more precise measure of the effectiveness of the training program. As shown in Figure 1, the levels are designed to measure different aspects of the training experience:

**Level I: Reaction** - measures the subjective views of the participants to the training, such as rating the overall quality of the training and the perceived effectiveness of the instructor, format, content, and delivery method (satisfaction data). It may also assess affective reactions such as level of satisfaction and enjoyment of the training, utility judgments, such as perceived relevance and practical value of the training for subsequent job performance;

**Level II: Learning** - measures the extent to which the participants acquired the principal facts, increased knowledge, attitudes, and/or increased skills (learned data);
Level III: Behaviour - measures behavioural changes that occurred as a result of the training and the knowledge transfer to improve job performance (performance data); and

Level IV: Results - measures the impact of the training program as it affects and contributes to the objectives of the organization (e.g., increased identification of suicidal individuals, reduced suicide incidents). The evaluation procedures are strategically designed to assist enhancing training program objectives and to modify training program delivery, and attend to organizational factors that may affect the transfer of knowledge to the job and related outcomes.

An important contribution of the framework is that the information obtained from the four levels of evaluation may be integrated to provide accurate and thorough evaluation of the training itself. Results from Level I (utility judgment) and Level II (learning) assessments can be combined with those from Level III (on the job behaviours) to gain a better understanding of the outcomes of the training. In that way, information can be synthesized regarding the extent to which the trainees perceived the training useful, learned the relevant skills, transferred this learning to improve performance on the job, and thereby contributed to the reduction in suicide incidents.

In addition to allowing the researcher to capture the breadth and depth of views that existed among TTC employees who had attended the suicide prevention programs, the evaluation framework also guided the researcher in developing the qualitative component of the evaluation, and assisted in the evaluation of the model as a conceptual framework guiding common methods for evaluating the effectiveness of gatekeeper programs.

**Structure of this Evaluation**

The structure of this evaluation is set out as follows:

- Chapter Two of this evaluation reviews current literature on subway-related suicides.
- Chapter Three examines program evaluations in general and more specifically suicide prevention program evaluations.
- Chapter Four provides a summary of the overall methodology used in this evaluation and describes the data analysis.
- Chapter Five presents the quantitative findings commencing with a description of the sample characteristics, and the changes in attitudes, knowledge, and skills acquisition.
- Chapter Six presents the findings from the qualitative component of the evaluation and presents them in terms of the Kirkpatrick Model, the theoretical framework that guided the evaluation.
- Chapter Seven presents the conclusions and discussion on suicide prevention evaluations.

Figure 1. The Kirkpatrick Model Pyramid (Winfrey, 1999).
CHAPTER TWO
SUBWAY SUICIDES: A REVIEW OF THE LITERATURE

Subway systems are a critical infrastructure element in urban centres internationally providing affordable, convenient, and rapid transportation to millions of urban communities in 135 countries (Gershon, Qureshi, & Barrera, 2005; Light Rail Transit Association, 2007). The subway structural environment, with its enormous volume of people maneuvering in transfer stations, crammed pathways, restricted narrow platforms, crowded stairs, escalators, and ramps provides numerous opportunities for adverse incidents to occur. Despite these challenges, actual subway-related injuries are remarkably low. Subways operate at high speed, capacity, and frequency and use routes that are independent from other forms of motorized traffic; thus, the structural environment of the subway system provides a distinct milieu within which incidents of injuries can occur (Gershon et al., 2008; Schwandl, 2007). One type of adverse event uniquely associated with subways is suicidal behaviour such as: either throwing oneself into the path of an oncoming train, lying on the tracks or on the “third rail” i.e., the electrified track. While subway-related suicide commonly refers to individuals who jump in front of moving subway trains, several authors have also described other types of suicidal behaviours, such as lying on the tracks in anticipation of a moving train and intentionally touching electrified rails (Chowdhury, Dutta, & Chowdhury, 2000; Guggenheim & Weisman, 1974; Rabban, Adler, Rosen, Blair, & Sheridan, 1997).

Subway-related suicides are a well documented problem associated with rapid transit systems worldwide (Baumert, Erazo, & Ladwig, 2006; Clarke & Poyner, 1994; Gaylord & Lester, 1994; Guggenheim & Weisman, 1972; Mishara, 1999; O’Donnell & Farmer, 1992; O’Grady & Griesi, 2002; Sohier & Sutton, 2002). An international survey of 50 subway systems conducted by O’Grady and Griesi (2002) found the annual number of subway-related suicides per system ranged from 0 to 120 with a reported mean of 10 suicides. Worldwide, the incident rate of subway-related suicides was 4.6 per 100 million passengers. European transit systems reported an annual incident rate of 4.9 per 100 million passengers, followed by North American transit systems which reported an annual incident rate of 4.4 per 100 million passengers, while Asian transit systems reported the lowest annual incident rate of 4.0 per 100 million passengers (O’Grady & Griesi, 2002). The highest annual number of subway-related suicides was reported
by European transit systems (mean 15; range 0 to 120), followed by Asian transit systems (mean 5.4; range 0.6 to 15), while North American transit systems reported the lowest number of subway-related suicides (mean 4.8; range 0 to 27). Recently, the Seoul Metro and Seoul Metropolitan Rapid Transit Corporation reported an increase of 30 percent in the number of subway-related suicides (Park, 2007). The number of subway-related suicides increased sharply from 21 in 2000 to 52 in 2003. The number of subway-related suicides which decreased to 34 in 2004, rose to 41 in 2005 and was expected to surpass 40 in 2007.

Ladwig and Baumert (2004) examined subway-related incidents in Munich from 1980-1999, and found a relatively stable incident rate which contrasts with the decreasing trend of all-cause suicide incidents in Munich. In the past decade, the Toronto Transit Commission (TTC) reports an annual mean of 15 subway-related suicides a year (range 8-22) (V. Cosentino, personal communication, May 22, 2008).

Contrary to the prevailing perception among the public and in view of the immense physical forces involved, subway-related attempted suicides are not uniformly lethal and the percentage of attempters who died varied from 20 to 80 percent, with the variance attributed to the operation and features of the subway structural environment (Cocks, 1987; Gaylord & Lester, 1994; Guggenheim & Weisman, 1972; Ladwig & Baumert, 2004; Mishara, 1999; O'Donnell & Farmer, 1994, 1992; O'Grady & Griesi, 2002).

Guggenheim and Weisman (1972) examining fatalities in the Boston subway from 1966 to 1972 found that only 32% of subway-related suicide attempts were fatal. In a later study of 100 patients injured by the London Underground trains from 1981-1986, 43% of the suicide attempts resulted in death (Cocks, 1987). Examining subway-related suicides in the Hong Kong subway system, Gaylord and Lester (1994) found that only 42% of the suicide attempts in the Hong Kong subway system resulted in fatalities.

In a review of the Montreal Coroner’s Office records from 1986 to 1996, Mishara (1999) found that less than 30% of the 323 subway-related suicide attempts were fatal. However, a third of the Montreal subway fatalities were classified as “undetermined” and further investigation revealed a history of suicidal behaviour among the decedents, which led researchers to conclude that the cases were misclassified (Mishara, 1999).

A similar attempt to chart suicidal behaviour over a period of one decade (1980-1990) in the Munich subway found that only 66% of subway-related suicide attempts resulted in fatalities.
Similarly, O’Grady and Griesi (2002) reported a 60% fatality rate in Toronto subway system. Ladwig and Baumert (2004); however, observed that the fatality rate was greater (83.3%) for suicide attempters over the age of 80 years old.

Injuries incurred by a person coming in contact with any part of the subway train are devastating. Among survivors, major extremity amputations are common. The individual and societal costs are enormous and result in “long-term social, functional, and financial consequences” (Guth, O’Neill, Pachter, & Diflo, 2006). It has been estimated that the direct cost to society of all train-related injuries in the USA may exceed $300 million per year (Guth et al., 2006).

In addition to the obvious loss of human life, suicides on the subway system have traumatic consequences for both transit employees and passengers; they place the public at risk due to train delays and overcrowding on narrow platforms, and have corporate and societal economic ramifications (Ladwig & Baumert, 2004). The high corporate costs associated with the an increasing number of subway-related attempted suicides in the Beijing subway system led the Beijing Subway Company, in 2002, to urge the Beijing Municipal People’s Congress for an early passage of legislation that would allow the prosecution of suicide attempters (Guihong, 2002).

According to the general-dispatcher of the Beijing Subway Company, most suicide attempters survived after receiving medical attention; however, each suicide attempt forced the closure of the subway line, and across the system delays resulted in overcrowding on platforms and a large number of passengers requesting ticket refunds and compensation. Additionally, the sudden braking of the train often led to injuries among the train passengers and damages to passengers’ personal property, for which the Beijing Subway Company was liable. Legislation, it was argued, would place legal responsibility on the suicide attempter and safeguard the rights of subway passengers to a rapid transit system free of interruptions and delays. “Society should not have to pay for an individual’s emotional behaviour” opined a Beijing lawyer.

While Beijing does not have any by-laws penalizing suicide attempters in the subway system, the Security Administration Punishment Act states that those who interfere with public traffic (buses, trolley buses, trains) would be prosecuted and face a 15-day jail sentence and a monetary fine. In other parts of China, both in Shanghai and Guangzhou, special legislation to deal with subway-related suicide attempters exists, and in Taipei and Singapore those who
attempt suicide on the subway system and interfere with “public order” are punished severely if they survive. Recently, a man who attempted suicide in the Singapore subway system was sentenced to six months in prison and was charged a large fine (Guihong, 2002).

The ensuing section of this chapter describes some of the individual characteristics, station structural features, and temporal patterns associated with suicidal behaviour in subway systems reported by Ratnayake, Links and Eynan (2007).

**Station Features: Passenger Volume, Station Structural Design, and Location**

The subway station location, structural design and passenger flow influences the number of suicide incidents. Several studies report higher numbers of subway-related suicide incidents in subway stations with greater passenger flow (Johnston & Waddell, 1984; Sonneck, Etzersdorfer, & Nagel-Keuss, 1994). Johnston and Waddell (1984) reported that transfer stations (i.e., Bloor-Yonge station, St. George station) have more subway-related suicide attempts than other stations within the Toronto Transit system. While the absolute number of suicide attempts at these high volume stations may be proportional to the passenger traffic, the ratio of attempts to passenger flow in a given period may be no higher than at any other station.

Similarly, Sonneck et al. (1994) found that the probability of suicide attempts increased with passenger volume in subway stations in Vienna; and while the number of subway-related suicide incidents was higher when compared to smaller subway stations, the ratio of suicide incidents to the daily number of passengers was lower at larger stations. In their examination of data from 1940 to 1989; however, Farmer, O’Donnell and Tranah (1991) found a higher number of suicide incidents in stations with higher passenger volume in the London Underground. The lowest number of suicide incidents was reported at terminal stations (V. Cosentino, personal communication, May 22, 2008; Sonneck et al, 1994).

Subway stations with greater proximity to psychiatric facilities were reported to have higher numbers of subway-related suicide incidents (Farmer et al., 1991; Johnston & Waddell, 1984; O’Donnell & Farmer, 1994). Johnston and Waddell (1984) found a greater number of suicide incidents in specific subway stations west of the Toronto downtown core. These subway stations (Spadina and Islington) were located in close proximity to two Toronto psychiatric facilities. O’Donnell and Farmer (1994) opined that subway stations with low passenger flow
and a high absolute risk of suicide incidents tended to be located in close proximity to a psychiatric facility.

Mishara (1999) found that 70% of subway-related suicides in the Montreal Metro system occurred in the station closest to the suicide decedent’s place of residence. In an earlier study, Farmer et al. (1991) concluded that stations that reported a high number of suicides had an increased proportion of victims with “no fixed abode”.

The structural design of subway stations contributes to the lethality of the suicide attempt. Subway stations in the London Underground that have water drainage pits beneath the rails were found to have a lower mortality rate as they provide a gap of space below the tracks which prevents suicide attempter from making bodily contact with the train (Coats & Walter, 1999). Between 1996 and 1997, a statistically significant reduction in the mortality rate among those individuals who attempted suicide in the London Underground stations that have drainage pits (44%) as compared to individuals who attempted suicide in stations that do not have drainage pits below the track level (76%) was noted. Elsewhere, O’Donnell and Farmer (1994) found that subway stations in the London Underground that have drainage pits had a significantly higher proportion of survivors (55%) as compared to subway station which do not have drainage pits (34%).

Most suicides in the London Underground (87.4%) and the Toronto subway system involved individuals jumping in front of the incoming train from the station’s platform. The majority of these incidents take place within 50 feet of the train’s entrance point into station. Reduced mortality is reported for individuals jumping off the platform beyond the train entrance point to the station. Few suicides occur on open track or in the tunnel.

Temporal Patterns: Day of the Week, Time of Day, and Seasonal Patterns

Subway-related suicides appear to be peaking on Mondays (Farmer et al., 1991; Sonneck et al., 1994) whilst Sundays are reported to have fewer suicide incidents compared to all other days of the week (Farmer et al., 1991; Ladwig & Baumert, 2004; O’Donnell & Farmer, 1994). However, Lin and Gill (2009) in their study of subway train-related fatalities in New York City found the greatest number of subway-related suicides occurred on Tuesday while the lowest number of suicide incidents occurred on Wednesday. Gender stratification; however, contoured the weekly distribution: the peak day for women was Monday (19%) whereas the peak day for
men was Tuesday (17.7%). The case fatality is also reported to be higher in the afternoon hours during the winter months (72%), with lethality peaking on Thursdays (78%). In contrast, survival was more likely in late night hours, on Saturdays, and in the summer months (Ladwig & Baumert, 2004).

Most subway-related suicide incidents tend to occur in late morning and early afternoon hours between the hours of 1100 and 1600 (Farmer et al, 1991; Gaylord & Lester, 1994; Guggenheim & Weisman, 1972; Johnston & Waddell, 1984; Ladwig & Baumert, 2004; Lin & Gill, 2009; Mishara, 1999; O’Donnell & Farmer, 1992, 1994). Few subway-related suicides were found to occur at night time and in all probability because most subway systems do not operate after mid-night. Guggenheim and Weisman (1972) reported that of the 50 suicide incidents they examined in Boston only 12 incidents occurred during night hours (18:00-24:00) all of which involved individuals with alcohol use disorders.

When Ladwig and Baumert (2004) stratified the data from their Munich sample of 306 subway-related suicides attempts by time of day and gender; however, different peak times emerged for males and females. Their findings suggest that significantly more women completed suicide between the hours of 9:00 -12:00 and 12:00 -18:00 while more men chose the evening and night hours (18:00-24:00) to end their lives in the subway system.

Extensive variation in the peak and low months for subway-related suicide incidents across reviewed studies is noted and no clear patterns related to months of the year are evident.

**Demographic Characteristics: Gender, Age, Race, Marital Status, Place of Residence**

Subway-related suicides constitute a complex and multifaceted phenomenon. A confluence of person-related factors impact subway suicide behaviour: demographics, mental illness, psycho-social factors, and proximal and distal stressful life events.

**Gender**

Most studies reviewed reported an increased number of subway-related suicides and suicide attempts among males as compared to females, in the range of 1.58:1 to 4.3:1 (M:F) (V. Cosentino, personal communication, May 22, 2008; Gaylord & Lester, 1994; Gershon et al, 2008; Lin & Gill, 2009; Mishara, 1999; O’Donnell & Farmer, 1992; Sonneck et al, 1994). When compared to gender distributions of all suicide in Canada, the United Kingdom and other
Western countries, however, it appears there is a smaller gender difference in subway-related suicides and suicide attempts (WHO, 2000).

Several researchers who have reported separate statistics for subway-related suicides and suicide attempts demonstrated a lower ratio of attempted suicides by males as compared to females (V. Cosentino, personal communication, May 22, 2008; Gaylord & Lester, 1994; Sonneck et al, 1994). In a study of 94 surviving attempters, O’Donnell, Arthur and Farmer(1994) found nearly equivalent numbers of males (n=49) and females (n=45). With respect to studies of overall suicide incidents, O’Donnell and Farmer (1994) reported a higher proportion of males as compared to females (1.78:1) in a large 2847 person sample of suicide incidents, from 1950-1990, in the London Underground. Ladwig and Baumert (2004) also found overall gender differences to be relatively small (1.15:1). Similarly, in an earlier study of subway-related suicides, Guggenheim and Weisman (1972) documented almost equal numbers of males and females in a sample of 50 consecutive suicide incidents in Boston. Data obtained from the TTC indicated that in the past decade (1998-2008) an overall small gender difference in suicide attempts was observed (M=148 vs. F= 108). Apart from these anomalies, it appears that males exhibit subway-related suicidal behaviour more often than females, a pattern which is consistent with trends for other violent methods (Denning, Conwell, King, & Cox, 2000).

Age

Data collected from the Chief Medical Examiner of New York City on 343 subway-related suicides, between 1990-2003, found that the suicide decedents were more likely to be between the ages of 25-54 (70%); 12.2% were younger than 24, and 28% were 25-34 years (Galea et al, 2006). Previous studies have found that the majority of individuals who exhibited subway-related suicidal behaviour were between the ages of 20 to 30 years old (Johnston & Waddell, 1984; Ladwig & Baumert, 2004; O’Donnell & Farmer, 1992;1994; Sonneck et al., 1994); however, Mishara (1999) reported most individuals (64%) who completed suicide in the Montreal subway system tended to be younger than 41 years of age. A more recent examination of data collected by the Office of the Chief Medical Examiner of New York City on subway-related fatalities (Jan. 1, 2003 and May 31, 2007) found the age range of the 111 suicide decedents was 14-85 with a mean age of 45.2 years (Lin and Gill 2009). Conversely, data from
the Tokyo system indicated a peak occurrence among individuals 51 to 60 years of age (O’Donnell & Farmer, 1992).

**Race/Ethnicity**

Of the studies reviewed, only two reported on the race/ethnicity of subway-related suicide decedents. Both studies examined subway-related suicides in New York City. Gershon et al. (2008) reported that of the 343 subway-related suicides that occurred in New York City during the years 1990-2003, 62.1 percent of suicide decedents were non-white (Black, Hispanic, Asian/other combined). Lin and Gill (2009) reported a similar race/ethnic distribution. Of the 111 subway-related suicides that occurred in New York City between January 1st, 2003 and May 31st, 2007 32% were Caucasian, 28% African-American, 28% Hispanic, and 11% Asian. The racial/ethnic distribution of the suicide decedents approximates the racial/ethnic distribution in New York City according to 2000 U.S. Census data (cited in Lin & Gill, 2009).

**Marital Status and Living Circumstances**

Few studies reported on the marital status and/or place of residence at the time of suicide (Guggenheim & Weisman, 1972; Mishara, 1999). Of those that did, studies indicate that the majority of suicide decedents were single and lived alone at the time of their suicide (Guggenheim & Weisman, 1972; Mishara, 1999). Guggenheim and Weisman (1972) found that of the 50 suicide incidents they studied in Boston, 62% were unmarried at the time of suicide. Congruent findings were reported by Mishara (1999) who found that of the 129 suicides in the Montreal sample, 60% were single, 24% were married, and 16% were widowed, separated or divorced (Mishara, 1999). In a sub-sample of 112 subway-related suicides from the Montreal study, nearly half (45%) lived alone, 28% lived with family, and 27% lived in mental health facilities at the time of their suicide. Mental health facilities included “psychiatric units of general hospitals, psychiatric hospitals, and halfway houses with supervised living and therapeutic activities” (B. L. Mishara, personal communication, 2005).

**Prevalence of Psychopathology**

Several studies found a high incident of diagnosed psychopathology in subway-related suicide decedents. Guggenheim and Weisman (1972) found that of the 50 cases of subway
suicides in Boston, 75% had received psychiatric care. The most common diagnoses at the time (due to labels used) were affective psychosis (37%), neuroses (32%), schizophrenia (18%) and organic psychoses (13%). Similarly, Johnston and Waddell (1984) reported that 61% of 119 subway-related suicides in Toronto had a previous psychiatric history of depression (37%), schizophrenia (13.7%), paranoia (6.8%) or an undetermined psychiatric illness (42.5 percent); however, less than 6% of the 119 suicide decedents examined were receiving in-patient treatment while less than 16% were receiving out-patient treatment at the time of suicide. High proportions of serious mental illnesses were also found in other studies (Cocks, 1987; O’Donnell, Farmer, & Catalan, 1996). Cocks (1987) reported that in a 100 person sample, 58% had received in-patient treatment at some time, 13% were in-patients at the time of their suicides and 2% had been discharged from a psychiatric facility within 48 hours of the incidents. Using data on 26 attempters, O’Donnell et al. (1996) found that 38% were psychiatric in-patients, 19% were receiving outpatient treatment, and 15% were being treated by general practitioners with antidepressants at the time of suicide. More recently, Mishara (1999) found that of 105 of the 129 subway-related suicides in his sample, 86% of individuals had been previously diagnosed with mental illness. The most common primary diagnoses were depression (50%), schizophrenia (25.5%) and non-specific psychosis (16.5%). Nearly 72% of the suicide decedents were receiving either in-patient treatment or psychotropic medications at the time of their suicide.

**Substance Use**

The role alcohol and/or drugs play in subway-related suicides and suicide attempts is difficult to ascertain as limited information is reported in the reviewed studies. Gershon et al. (2008) reported that among 343 subway-related suicides in New York City, one or more illicit drugs were detected in 21.3% of the suicide decedents, and alcohol was present in 17.8%. Lin and Gill (2009) reported that among the 111 subway-related suicides they have examined toxicology reports indicated cocaine and/or benzoylecgonine and/or alcohol were detected in 14% while. Antidepressant medications were detected in 21% of the suicide decedents. A positive toxicology report for both illicit drugs and alcohol was detected in 31.5% of subway suicide decedents. In a sub-sample of 78 suicides for which the information was available, Mishara (1999) reported that autopsy blood analyses detected alcohol present in 25% of the subway suicide decedents, and 24% had drugs present in their bloodstream; however, the levels
of alcohol and drugs were not high and included traces of psychotropic medications. In an earlier study, Cocks (1987) reported that in a sample of 100 decedents in the London Underground, where suicide was suspected, the clinical and post-mortem findings suggested that in approximately three quarters of the sample only 9% of the decedents had consumed substantial amounts of alcohol. Witnesses’ accounts indicated that 6% of those who had completed suicide had been “visibly drunk” prior to their suicides. Elsewhere, Guggenheim and Weisman (1972) reported that 18% and 24% of the 50 suicide incidents they studied suffered from acute and chronic alcoholism, respectively.

**Previous History of Suicide Attempts**

A history of previous suicide attempts was found to be common among subway suicide decedents (Cocks, 1987; Johnston & Waddell, 1984; Mishara, 1999). Mishara (1999) found in a sub-sample of 88 suicides for which data was available, two-thirds (66%) of Montreal subway-related suicide decedents had a history of suicidal behaviour: 20.5% had one previous attempt, 22.7% had two previous attempts, and 22.7% had more than two previous attempts. An additional salient finding was that 9% of subway suicide decedents had made a previous suicide attempt in the subway system. Johnston and Waddell (1984) reported that in a sub-sample of 73 subway-related suicides, 40% made an unspecified number of previous attempts. Similar findings were reported by Cocks (1987).

**Subsequent Suicide Attempts**

An examination by O’Donnell et al. (1994) of subsequent mortality records of 94 survivors whose index attempts occurred between 1977 and 1979, found that over a 10-year period, approximately 10% of the survivors of London Underground suicide attempts subsequently die by suicide, a percentage comparable to rates of survivors of other suicide attempt methods. All subsequent suicides occurred within a period of 3 years. In an earlier study, Guggenheim and Weisman’s (1972) data on the subsequent suicidal behaviour of 33 survivors of subway-related suicide attempts in the Boston subway system indicated that at least 21% of the survivors attempted or die by suicide during 5.5 years of follow-up.
Communicating Suicide Intent

Several studies reported that individuals may have indicated a desire to attempt suicide prior to the suicide incident in the subway; however, these individuals often may not have indicated their intention to use the subway as means to end their life. In Mishara’s (1999) study, information on the past behaviours of only 100 of 129 subway-related suicides was available. Of those, 81% expressed a desire to end their life and most had articulated their intent several times prior to their suicide. In a sub-sample of 50 individuals for which the information was available, 26 had expressed a desire to kill themselves on the day of their suicides, and 21% of the original sample left a suicide note. In an earlier study, Cocks (1987) showed that 15% of 100 patients injured in the London Underground expressed their suicidal intent during the 24 hours preceding their suicide attempt. None of these individuals; however, identified the subway as their planned means of suicide.

Elsewhere, O’Donnell, Farmer and Catalan (1993) found in a sample of 242 probable suicides and 138 probable suicide attempts in the London Underground, only 15% left suicide notes. Analysis of 37 of the 42 available notes revealed that only five individuals had stated they intended to use the subway as means of suicide and of those five, four of their notes mentioned a specific station for the planned attempt.

Conclusions

Several reasons for the use of the subway as a means to end one’s life have been proposed in the literature reviewed. Distressed individuals intent on ending their life may lack the privacy to use other methods, or may feel they do not have to be concerned about having last-minute doubts, and/or may perceive the method to be highly lethal (Clarke & Poyner, 1994; Guggenheim & Weisman, 1974; O’Donnell et al., 1996). Case studies indicate that the chosen method of suicide is often determined as the culmination of a preconceived premeditated plan rather than as a hasty impulsive decision (Guggenheim & Weisman, 1974). Others have argued that subway suicide attempters are highly impulsive and their act is characterized by an extremely high level of suicidal intent (Ladwig & Baumert, 2004; O’Donnell et al., 1996).

Rates within cities appear to be independent of national rates of all suicide methods. Using mean annual rates previously reported elsewhere, Lester (1995) asserted that there was no
association between subway rates in 17 cities around the world and each city’s corresponding national suicide rate in 1980.

The outcome of a suicide attempt is not only the result of the lethality of the suicide means but also a result of a complex interplay between personal and environmental factors. Suicides occur more often at subway stations located in close proximity to psychiatric facilities or the decedent’s place or residence. The physical design of the subway station i.e., presence of drainage pits or barriers on platform, and the location on the platform from which the individual jumps into the pathway of the incoming train determines the lethality of the incident. The epidemiological characteristics of those who choose the subway system as a method to kill themselves are markedly distinct from characteristics of other individuals in the general population who die by suicide (Ladwig & Baumert, 2004; Lester, 1995). Violent suicide methods follow a more pronounced seasonality than softer methods, with a peak in spring and a low in late autumn (Hakko, Rasanen, & Tiihonen, 1998; Preti & Miotto, 1998). A similar result was found for the weekly distribution with violent suicides peaking in the beginning of the week (Angermeyer & Massing, 1985). This was also observed in railway-related suicides (Schmidtke, 1994). Subway suicides do not follow long established circadian times and seasonal patterns. Data from reviewed studies does not support a clear high risk day of the week. The studies reported remarkably similar findings, giving evidence that individuals who attempt suicide on the subway may present as a unique subgroup of suicidal individuals.

Saving lives, averting devastating physical injuries, and preventing the psychological trauma to train operators and bystanders by preventing subway-related suicidal behaviour is imperative. Preventing subway-related suicidal behaviour can consist of several diverse strategies:

- Educating clinicians about the risk factors associated with subway suicide;
- Restrict media reporting of subway suicides;
- Structural modification to existing subway stations: 1) erecting safety barriers on at least a portion of the subway platform edge i.e., the first 50 feet from the trains’ entrance point to the station; 2) adding drainage pits; and 3) installing surveillance devices.
- Reducing the traveling speed of the train as it approaches the subway station to ease the stopping of the train in cases of suicidal behaviour;
• Implementing gatekeeper suicide prevention programs which would increase employees’ vigilance and willingness to intervene.

Suicide is complex and multifaceted issue that involves psychological, social, economic, genetic, cultural, and environmental factors, and thus no single strategy is likely to have the potential to reduce the incidents of subway-related suicides. Coalescing diverse suicide prevention strategies to address various suicide risk factors is required to achieve reduction in subway suicide rates. The available evidence thus far suggests that among most promising strategies are:

“Programs which focus on enhancing the skills of community, organizational and institutional gatekeepers (including, for example, workplaces) can increase public awareness about suicide, encourage help seeking, and improve identification and referral of those at risk for suicide.” (Beautrais, 2006, p. 6).
CHAPTER THREE

PROGRAM EVALUATION AND THEORETICAL APPROACHES

Since the major purpose of this study was to evaluate a suicide prevention training program, the first part of this chapter provides an overview of program evaluation models and approaches applied to mental health. The latter part of the chapter provides an overview of gatekeeper suicide prevention program evaluations.

Program Evaluation

The field of health promotion is under increasing pressure to demonstrate that programs are worthwhile, effective and efficient. The need to evaluate outcomes, processes and structures is identified in reviews of training programs in health and mental health settings (Cook & Shadish, 1986; Goldstein, 1977). Patton (1997, p. 23) defined program evaluation as “the systematic collection of information about the activities, characteristics, and outcomes of programs to make judgments about the program, improve program effectiveness, and/or inform decisions about future program development.” The term "program" may include any organized action such as a media campaign, service provision, educational services, public policy, research project, etc. (CDC, 1999). Rowntree (1982, 1992) proposed that although the terms evaluation and assessment are often used interchangeably, the terms are not synonymous as they refer to different levels of investigation. While evaluation is concerned with the macro or holistic level of the training and considers the context of the training and all the concomitant factors, assessment can be seen as a measure of the trainee’s learning and is one of the elements that is incorporated into the evaluation. The four points of focus for health promotion evaluation have been identified by Schulburg (1982) as the assessment of: 1. effort, 2. effectiveness, 3. adequacy, and 4. efficiency. Effort relates to the evaluation process, effectiveness to the evaluation outcome, and adequacy and efficiency to managerial and structural measures of accountability, respectively. Cahn (1975) defined outcome as the degrees to which the trainee and the services provided by the trainee are altered as a consequence of the training. Process, on the other hand, is defined in terms of how appropriately and completely methods of training are applied, while structure refers to the curriculum and learning materials used in the training program (Cahn, 1975).
Rossi, Lipsey and Freeman (2004) divided the dimensions of program evaluation into five main categories: needs assessment, program theory, process analysis, impact analysis, and cost-benefit and cost-effectiveness analysis. The needs assessment determines the nature of the needs the program is intended to address. The program theory is the formal description of the program’s concept and design and examines the salient components of the program and the anticipated short- and long-term effects. This is also referred to as a logic model or impact pathway. “Logic models are descriptive” (Patton, 2002, p. 163), they provide in graphic form the chain results, connecting activities to the expected initial outputs, and intermediate and final outcomes. It provides a basis for developing the performance measurement and evaluation strategies (Patton, 2002, pp. 162-163).

The analysis of the program theory also appraises the program organization and how it would lead to the desired outcomes. Additionally, the program theory may also reveal unintended or unforeseen, positive or negative, consequences of a program. The program theory guides the hypotheses to be tested for impact evaluation and determines the causal effects of the program, whereas process analysis looks beyond the theory of what the program is supposed to deliver and instead evaluates how the program is being implemented. Finally, cost-benefit or cost-effectiveness analysis assesses the efficiency of the program in terms of returns on investments (ROI) for the stakeholders.

Evaluation Approaches

Effective evaluation requires a systematic, eclectic, multi-method approach where methods and procedures appropriately fit the specific program evaluation objectives. Evaluation designs describe the key features and procedures to be followed in conducting an evaluation. Fitz-Gibbon and Morris (1987, p. 9) have defined an evaluation design as “a plan which dictates when and from whom measurements will be gathered during the course of an evaluation;” hence, the program evaluation design ought to accommodate the complexity of program activities and meet the needs of diverse stakeholders. As a result, the use of multiple methods to evaluate program efforts is often the most comprehensive approach in a synthesis of research methods which includes a fusion of the two meta-methodologies: quantitative and qualitative approaches.

Although there appears to be consensus among researchers on the definition of the term evaluation, a review of program evaluation literature reveals diversified views on the meaning of
conducting program evaluation. The diversity is particularly evident in different orientations that exist toward the utilization of evaluation methodology and evaluation findings. Some researchers (e.g. Patton, 1978; Stufflebeam, 1971) emphasize the informational needs of organizations or the numerous uses found in various decision making contexts such as facilitating improvements, generating knowledge, and making overall judgments. This heterogeneity extends to methodological approaches. Some evaluations attempt to simulate quasi-experimental design while others employ system analysis techniques for data collection. A review of recent health professions literature demonstrates an increased interest in qualitative evaluation methods which represents a paradigm shift from the more traditional, conventional, quantitative evaluation methodologies (Cohen & Crabtree, 2008).

Qualitative approaches differ from quantitative approaches in the philosophical assumptions about the goals of the evaluation inquiry. While qualitative evaluation evolved from a phenomenological philosophical perspective that favours an interpretive approach to social inquiry and emphasizes the particular over the universal, quantitative orientation developed from a positivistic philosophy of science with its search for parsimony (Pearsol, 1985, p. 132). The goal of positivist perspective is to generate objective knowledge independent of historical, societal or individual conditions. The positivistic paradigm relies on a “hypothetico-deductive method” (Whitly & Crawford, 2005), where data is collected to test a theoretically guided, predefined hypothesis. Knowledge is transmitted and confirmed over time as theories and explanations and/or as predictions (Farley, 1982). “Applied” positivism thus adheres to scientific method and quantifies knowledge: propositions are transformed into hypotheses that are either accepted or rejected based on a predetermined level of probability; hence, proof is based on statistical models that allow researchers to make assumptions about causation or correlation of observed facts (Pearsol, 1985).

Explanations in social sciences therefore, are based on the recognition and understanding that knowledge is formed in different contexts and is influenced by cultural meanings and social actions. The goal of the qualitative inquiry is to seek causes and facts from the etic or Weltanschauung (world view) perspective (Osborne, 1977). The emphases are on the subjective meanings and experience in all of their complexities and multiplicities. The participant’s perspective has priority over the researcher’s preconceived framework. The findings are based on the participant’s and the researcher’s interpretations of events. An aphorism used by many
qualitative researchers is that the participant is the expert, whereas the researcher is the layperson—a reversal of the traditional biomedical epistemological assumptions (Whitley & Crawford, 2005).

Contrary to qualitative evaluations that result in outcomes or conclusions that inform action, conventional quantitative evaluations result in outcomes that signal a category of conclusions. This scientific paradigm approach includes both objective-based and experimentally-oriented evaluations. The objective-oriented approach focuses attention on program impact or outcomes and it has blended well with the experimental design model of research, often using experimental methodology to study impact.

While conventional evaluation strategy is contingent on monitoring input and measuring output, qualitative evaluation provides explanations that illuminate motives, values and provide insight into the context of meaning within which an activity is or will be performed; thus, “qualitative evaluations are oriented towards exploration, discovery, and inductive logic” (Patton, 1987, p. 15). They emphasize “the shape of relations between reasons and events, not between causes and effects” (Niklas, 1982, p. 663). Consequently, the emphasis is on understanding rather than quasi-experimental explanation. Some evaluation researchers had suggested that combining evaluation approaches might result in a better evaluation than utilizing either a qualitative or quantitative approach (Cook & Reichardt, 1979; Madey, 1982; Montagne, 1982).

Comprehensive program evaluation requires multi-strand investigations allowing for the use of the best of both quantitative and qualitative methods in response to the particular evaluation questions of interest in an evaluation.

The underlying logic of mixing methods is that neither quantitative nor qualitative methods are sufficient in themselves to capture the intricacies of a program. Using more than one approach allows the evaluator to combine strengths and correct some of the deficiencies of any one source of data (Greene, Caraccelli, & Graham, 1989; Patton, 1987; Rossman & Wilson, 1985). When used in combination, both quantitative and qualitative approaches add complexity to the evaluation design and apply the advantages of both the quantitative and qualitative approaches (Creswell, 1994; Gogolin & Swartz, 1992; Greene et al., 1989; Grinnell, 1997; Johnson & Onwuegbuzie, 2004; Kushman, 1992; Madey, 1982; Patton, 1987; Rossman & Wilson, 1985).
There are extensive arguments on how mixed methods allow for improvement of the accuracy of conclusions by relying on data from more than one method. Greene et al. (1989), for example, proposed five major rationales for the integration of mixed-model approaches in program evaluations:

- **Triangulation** - it allows for the examination of data collected through more than one method to seek convergence in the findings.
- **Complementarity** - it permits elaboration, enhancement, clarification, and illustration of results from one method with results from the other method. It provides richness in detail expanding understanding of the participants’ reactions to the program and perceived relevance to their work. It explores interconnected and/or distinct aspects of the data.
- **Initiation** - it facilitates the discovery of inconsistencies and contradictions in the data that could possibly lead to a substantial alteration and re-framing of the research question.
- **Development** - it allows for augmentation by using the findings from one method to help inform and/or develop the other method which enhances the validity of the constructs by capitalizing on the inherent strengths of both methods.
- **Expansion** - it increases the breadth and depth of the results and interpretation by analyzing them from different perspectives.

### Evaluating Workplace Education Programs

The program planning literature (Kellaghan & Madaus, 2000; Patton, 2002; Rando & Lenz, 1994; Weston, McAlpine, & Bordonaro, 1995) distinguishes between different types of workplace education program evaluations. There are two general categories in which to group evaluation processes: formative or process evaluation and summative or outcome evaluation.

### Formative Evaluation (Process Evaluation)

Formative evaluation, as implicated by its name, involves those evaluation processes that are generally conducted during the beginning or the middle stages or are conducted during the development phase of a program operation. These processes assist in identification of problem areas that can be addressed and modified while changes are possible and productive for the workplace education program. Formative evaluations are conducted with the intent to improve
programs and learning activities (Rando & Lenz, 1994). Formative evaluation focuses on the iterative process of development and testing program elements to ensure the program delivers meaningful outcomes (Government Accountability Office, 2005) to the recipients or stakeholders. Frequently, these evaluations are used to determine if all stakeholders have shared goals and a common understanding of the program components, processes, definitions, and directions. Additionally, these evaluations are used to determine whether the learning processes and methods utilized are sufficient in accomplishing the goals of all persons involved in the workplace education program. The evaluation tools or instruments that are used when conducting formative evaluation include: classroom evaluations, needs assessments, curriculum, supervisory staff and employee interviews. The information collected is meant to inform change. Its purpose is to validate or ensure that the goals of the instruction are being achieved, and to improve the instruction by providing intermittent feedback to all workplace education partners (Weston et al., 1995).

**Summative Evaluation (Outcome Evaluation)**

In contrast to formative evaluations, summative evaluations take place at the end of a program’s operation. They are designed to provide information on the program efficacy and to determine if the program and learning activities, usually in the aggregate, worked in terms of overall goals and objectives. The focus of summative evaluation is on outputs and outcomes and provides short-term effectiveness or long-term impact information. These evaluation measures (pre- and post- program data) usually assess the basic skills of the participants, the transfer of the skills to the job and workplace, the adult learner’s beliefs about basic skills and education, and the transfer of training in terms of measurable outcomes identified in the initial stages of the workplace education program. Summative evaluations use the following measurement instruments: pre and post employees’ questionnaires (related to basic skills practices); pre and post interviews with supervisory staff and employees, and subsequent performance ratings (identified as measurable outcomes) (Government Accountability Office, 2005; Weston et al., 1995).

Any program evaluation, formative or summative, is a multi-stage process involving setting of objectives, the execution of strategies, the collection of data, and the appraisal of the relative success or failure of a given program. These methodologies are critical aspects to
developing an evaluation plan that ultimately provides outcome and feedback to stakeholders. The goals of gatekeeper suicide prevention programs are to increase awareness about suicide prevention and to help “emergent” and “designated” gatekeepers develop competencies and intervention skills to help persons at risk of suicide. Summative program evaluations have been deemed as critically needed in the area of suicide prevention (Kaleveld & English, 2005) to evaluate the salient components of gatekeeper programs. The following section of this chapter begins with a brief discussion on the impetus for the development of suicide prevention training programs and examines the need for program evaluations in suicidology. The latter part of the chapter provides a review and a critique of existing gatekeeper suicide prevention summative evaluations.

Suicide Prevention Program Evaluations

Suicide is a complex multifaceted behaviour resulting from an intricate interaction of neurobiological, genetic, psychological, social, cultural, and environmental risk and protective factors (Mann et al., 2005). The study of suicide has become a largely interdisciplinary effort with theoretical and research contributions from the fields of psychiatry, psychology, sociology, philosophy, social work, medicine, nursing, biology, law, and theology. In addition to the increased scope of study in the field, the second half of the last century saw significant movement towards broad-based efforts in the prevention and treatment of suicidal behaviours. Part of these efforts is related to the development of training programs to meet perceived, manifest, and expressed needs in training in suicide prevention.

The impetus for training in suicide prevention stemmed from a number of factors which support the proposition that suicidal behaviour is: multidimensional (Leenaars, 1995); often preventable (Leenaars, Maltberger, & Neimayer, 1995; Morgan, 1981; Murphy, 1983; Shneidman, 1970; 1985); ambivalence about living or dying is frequently present (Shaffer, Garland, & Whittle, 1988; Shneidman, 1970; 1985); suicidal ideation is generally transient, and suicidal intent is usually communicated and warning signs are often evident (Shneidman, 1985).

Concomitant to the increased knowledge regarding suicidal behaviour, a shift in the responsibility for suicide prevention emerged during the community psychology movement of the 1960’s (Shneidman, 1970; 1985). Caplan’s (1964) model of community health and the analogous formulations proposed by Shneidman (1970) contributed to suicide prevention being
considered as a societal responsibility rather than limited to specialized professionals. The development of suicide prevention initiatives aimed at involving community members as gatekeepers originated from the realization that no one group in the community can take exclusive responsibility for suicide prevention (Potter, Powell, & Kachur, 1995; Snyder, 1971).

Subsequently, emergent theory based on research findings in suicidology (Beck, Kovacs, & Weissman, 1975; Maltsberger, 1988; Shneidman, 1970, 1985) provided the foundation for the development of suicide prevention training programs with educational components designed to enhance caregiver intervention competencies by providing factual information about suicide, increasing attitudinal awareness, and developing risk assessment and intervention skills (Lang, Ramsay, Tanney, & Tierney, 1989).

Several gatekeeper suicide prevention programs were developed over the years, to raise public awareness about suicide and suicide prevention. The programs developed vary in a number of ways:

- Targeted population: some are appropriate for everyone in the community while others are designed for one specific stakeholder population, such as school personnel or law enforcement personnel;
- Educational strategies: programs may include classroom or lecture-style information dissemination, small group discussion, use of videos with case studies, and/or participant scenario role plays; however, many programs include combinations of these educational strategies.
- Format: programs vary in length from very brief educational sessions to multiple day trainings.
- Content: a few programs combine gatekeeper trainings with screenings.

Although many suicide prevention programs have been developed and implemented, only a few programs have been formally evaluated for their effectiveness. Evaluation efforts in the field of suicide prevention training have been very limited and there is relatively little knowledge about the types of strategic or program-level interventions that successfully prevent suicide (Beautrais, 2005; Breton et al., 2002; CASP, 2004; CDC, 1994; Goldney, 1998a; 1998b; Mann et al., 2005; US Department of Health and Human Services, 2001; Tierney, 1994; Stein & Lambert, 1984). There is a lack of evidence-based approaches to suicide prevention programs. No comprehensive, standardized models of evaluation have been developed whereby training
and the effects of various forms of training in suicide prevention can be measured and compared across services (Stein & Lambert, 1984).

National Registry of Evidence-based Programs and Practices

In 2003, in response to Objective 10.3 of the United States National Strategy for Suicide Prevention (NSSP), which called for the development of a registry of prevention activities with demonstrated effectiveness for suicide and suicide behaviour, the Suicide Prevention Resource Center (SPRC) and the American Foundation for Suicide Prevention (AFSP) developed an online Best Practices Registry (BPR). The registry, which is part of the National Evidence-Based Practices Project (EBPP), was launched in the spring of 2004. The Best Practices Registry for Suicide Prevention (BPR) was developed to identify, review and disseminate information about best practices that address specific objectives of the National Strategy for Suicide (NSSP).

Programs listed on the BPR are categorized into three sections:

- **Section I: Evidence-Based Programs**: This section contains interventions that have undergone rigorous evaluation and have demonstrated positive outcomes;
- **Section II: Expert and Consensus Statements**: This section lists statements that summarize the current knowledge in the field and provide "best practice" recommendations to guide program and policy development.
- **Section III: Adherence to Programmatic and Safety Standards**: This section contains suicide prevention programs and practices including awareness materials, educational and training programs, protocols, and policies that have been implemented in specific settings. These programs address specific objectives of the NSSP and their content has been reviewed for accuracy, likelihood of meeting objectives, and adherence to program design standards.

The three sections do not represent levels of effectiveness, but rather include different types of programs and practices reviewed according to specific criteria developed for that section. The registry is located on the Suicide Prevention Resource Center website (www.sprc.org) and the programs listed were reviewed by experts using ten well-established standards for research quality. The criteria items were: theory, intervention fidelity, design,
attrition, outcome measures, psychometric properties analysis, safety, integrity, utility, and other plausible threats to validity. The following twelve programs were reviewed and classified as evidence-based by SPRC/AFSP.

- Community-Based Programs
  - United States Air Force Suicide Prevention Program*
  - Reduced Analgesic Packaging

- Emergency-Room Programs
  - ER Means Restriction Education for Parents
  - Emergency Room Intervention for Adolescent Females*

- Primary Care
  - PROSPECT (Prevention of Suicide in Primary Care Elderly: Collaborative Trial)*

- School-Based Programs
  - C-Care/CAST*
  - CARE (Care, Assess, Respond, Empower)
  - CAST (Coping And Support Training)
  - Columbia University TeenScreen*
  - Lifelines
  - Reconnecting Youth
  - SOS Signs of Suicide*
  - American Indian Life Skills Development/Zuni Life Skills Development*

- Service Delivery
  - Psychotherapy in the Home

An additional eight programs are listed under Section III (Adherence to Standards) of the SPRC/AFSP Best Practices Registry (BPR). Eight programs are considered gatekeeper suicide prevention programs for high schools, higher institutions of learning, and the community at large. The programs are:

- ASIST (Applied Suicide Intervention Skills Training)
- safeTALK
- LEADS for Youth (Linking Education and Awareness of Depression and Suicide)
- Response: A Comprehensive High-School Based Suicide Awareness Program
- Campus Connect: A Suicide Prevention Training for Gatekeepers
- Frameworks Suicide Prevention Program
- High School Gatekeeper Curriculum
- QPR Gatekeeper Training (Question, Persuade, Refer).

While these programs address specific goals of the National Strategy for Suicide Prevention and have been reviewed by a panel of experts and found to meet standards of accuracy, safety, and programmatic guidelines, the programs were not reviewed for evidence of effectiveness.

Since 2005, the National Registry of Evidence-based Programs and Practices have been managed by the Substance Abuse and Mental Health Services Administration (SAMHSA), an agency of the United States Department of Health and Human Resources (http://www.nerp.samhsa.gov/listofprograms.asp downloaded 3/7/09). The NREPP listed programs were required to utilize superior evaluation methods to demonstrate the strong causal link between the programs and appropriate outcomes.

Eleven suicide prevention and treatment programs are listed in the NREPP registry; eight programs are suicide prevention programs and three are treatment programs. The programs listed in the BPR above and designated with an asterisk (*) are also listed in the NREPP. Additionally, the NERPP lists three treatment programs:
- Cognitive Behavioural Therapy
- Dialectical Behavioural Therapy
- Multisystemic Therapy with Psychiatric Support (MST-Psychiatric).

The registry of eleven programs represents an initial step in the collection and promotion of evidence-based suicide prevention programs.

**Evaluations of Suicide Prevention Strategies**

A recent systematic review by Mann et al. (2005) examined the efficacy of suicide prevention strategies worldwide. The review identified five major areas of prevention: education and awareness programs for the general public, professionals and gatekeepers; screening methods for high-risk populations; treatment of psychopathology; restricting access to lethal
means; and media reporting guidelines for suicide. Heterogeneity in study methodology and populations circumscribed the authors’ ability to conduct a formal meta-analysis of the results; thus, a narrative synthesis was presented. The most promising interventions identified by Mann et al. (2005) were physicians’ education, means restrictions, and gatekeepers’ education. Gatekeeper education was most effective “where the roles of the gatekeepers are formalized and pathways to treatment are readily available” (Mann et al., 2005, p. 2071).

To date, the systemic evaluation of gatekeeper suicide prevention programs’ impact on suicide rates has been limited to multilevel programs in institutional settings such as military forces. The U.S. Air Force Suicide Prevention Program implemented a comprehensive, institution-wide intervention that focuses on enhancing protective factors and decreasing risk factors for suicide (N=5,260,292). Major goals of the program included:

1. Promoting awareness of the range of risk factors related to suicide;
2. Educating the community regarding available mental-health services; and,
3. Reducing the stigma related to help-seeking behaviors.

The program was evaluated using an interrupted time series design (Knox, Litts, Talcott, Feig, & Caine, 2003). Rates of suicide deaths, as well as other violence-related statistics, were examined for six years prior to and after program implementation. Analysis of post-implementation rates (1997-2002) revealed a 40 percent risk reduction for suicide (Mann et al., 2005, p. 2070). Similarly, subsequent to the implementation of a gatekeeper program in the Norwegian Army, a 33 percent decline in annual suicide rates was observed (Mann et al., 2005, p. 2070).

**Gatekeeper Program Evaluations**

Despite the widespread implementation and recommendation for gatekeeper training, there is a lack of consensus as to its effectiveness. The Center for Disease Control (CDC) in Atlanta recommended that three kinds of evaluations should be considered when evaluating gatekeeper programs: 1) assessment of the degree to which these programs have sensitized gatekeepers to their role in identifying and appropriately helping those who might be at risk of suicidal behaviour; 2) assessment of the degree to which these programs result in appropriate identification and referral of suicidal individuals; and 3) assessment of the impact of these programs on suicide and suicide attempts (www.epo.cdc.gov/wonder/prevguide/p0000024.asp.
The CDC acknowledged that assessing the impact of gatekeeper programs on rates of suicides in the community is extremely difficult. Intermediate outcomes that are easier to assess include changes in people’s knowledge of suicide warning signs, their attitudes toward seeking or providing help, and the referral of high-risk individuals to treatment.

The following section of this chapter reviews the evidence for the effectiveness of selected gatekeeper programs and examines their methodological strengths and limitations. Similar to a more recent literature search conducted by Isaac et al. (2009), a search for English-language articles was carried out using OVIDSP. The search included the following databases: MEDLINE (articles from 1950 to April week 4 2008) and PsycINFO (articles from 1960 to April week 4 2008). The key words searched were suicide, suicide prevention, and gatekeeper and combinations of suicide and gatekeeper, as well as suicide prevention and gatekeeper. The key words were present in the title, abstract, or both. The search for suicide and gatekeeper key words yielded 29 articles of which three were included in this review while the search suicide prevention and gatekeeper key words produced 25 articles and the same three articles were kept for inclusion in this review. As articles of interest for this literature review were those with a focus on program evaluation, further searches were conducted using the following key words suicide prevention, gatekeeper, program evaluation and combinations of these key words. The search for suicide prevention, gatekeeper, and program evaluation in PsycINFO yielded four articles while MEDLINE produced six articles of which three were kept and included in the review (Capp, Dean, & Lambert, 2001; Cross, Matthieu, Cerel, & Knox, 2007; Matthieu, Cross, Batres, Flora, & Knox, 2008). An additional search in MEDLINE using the terms suicide intervention, training, and evaluation yielded five articles of which two were included in the review (Neimeyer & Pfeiffer, 1994; Tierney, 1994). Gatekeeper program evaluations included in this review were suicide prevention programs implemented in the workplace; however, training provided to personnel working in primary care or emergency department settings were excluded from this review as such groups provide a more direct clinical service. Of the pertinent articles found, the references were scanned to further find articles discussing gatekeeper program evaluations. The internet was also searched for reports of the Applied Suicide Intervention Skill Training (ASIST), safeTALK, and suicideAWARE program evaluations. Seventeen ASIST evaluation reports were found; a number comparable to the fifteen ASIST evaluations reviewed and reported in the Choose Life review of the international literature. A total twenty-three
articles were found to be pertinent and thus included and reviewed. There are a number of limitations of this review to be acknowledged. First, although a comprehensive literature search was conducted, it was not by any means exhaustive. There is always the possibility for program evaluations to be missed. The second, the review excluded suicide prevention program evaluations which targeted specific groups in the community, i.e., adolescents, crisis lines, health-care workers. Third, the reviewed evaluations were of standardized and widely disseminated and implemented programs. Fourth, excluded from this review are programs with limited content and narrow scope, i.e., programs implemented with the intent to raise suicide awareness and which did not address attitudes or sought to enhance the intervention knowledge and skills of program attendees.

**Evaluation of Suicide Prevention Gatekeeper Programs in the Workplace**

One feasible site in which to train community gatekeepers is the workplace. Employee in-service education and training are standard aspects of workforce infrastructure and bring together employees with similar occupational roles to improve workplace safety and provide continuing training. Cross et al. (2007) examined the immediate outcomes of a 1-hour community gatekeeper suicide prevention training provided to 76 non-clinical employees (e.g., administrators, research coordinators and secretarial support) in a university hospital workplace setting. The evaluation used a repeated measure design with survey questionnaires completed by participants immediately pre- and post-training. A follow-up skills assessment was conducted six weeks after completing the training. Participants completed survey questionnaires which assessed their factual knowledge about suicide and perceived ability to identify and intervene with an individual at risk for suicide. Participants’ role-play experiences and satisfaction with the gatekeeper program were also evaluated. Cross et al. (2007) found the training resulted in positive immediate outcomes. Participants’ knowledge about suicide warning signs and how to intervene with someone who is suicidal were significantly enhanced after the training. Participants also felt more efficacious intervening with someone they thought might be at risk for suicide. Observational ratings of participants’ gatekeeper skills during a standardized role-playing immediately after the training and six weeks later indicated that 55% of participants in the brief gatekeeper training had acquired acceptable skills. The training was rated positively with 89.5% of the participants indicating they were satisfied with the training, 90.8% of
participants rated the training as valuable experience, and 94.8% stated they would recommend the training to others.

Elsewhere, Matthieu et al. (2008) reported on a gatekeeper suicide prevention program evaluation conducted with a national sample of community-based counselling centre employees from the U.S. Department of Veteran Affairs. Participants included clinical (i.e., psychologist, social worker, etc.) and non-clinical (i.e., administrative staff and community outreach workers) personnel. The gatekeeper program consisted of the QPR (Question, Persuade and Refer) workshop. The QPR is a brief; one-hour standardized community gatekeeper suicide prevention training created by Paul Quinnett (2007) and was provided by a certified QPR Institute. The training also included a gatekeeper skills practice activity offered in a three-person peer group format immediately after the large group presentation. The practice was a five to seven minute, standardized role-playing dialogue, which required the use of three gatekeeper skills. The evaluation consisted of a repeated measure design of pre- and post-training questionnaires. The pre- and post-training questionnaires assessed self-efficacy and factual knowledge. The post-training questionnaires also assessed satisfaction with the training and evaluated the previously described role-playing skills. Of the 760 trainees 79.2% (602) participated in the evaluation; 428 clinical and 174 non-clinical personnel. Both clinical and non-clinical cohorts had significantly higher scores on all outcome measures with a greater mean score and Cohen’s d effect size noted in the non-clinical cohort when compared to the clinical cohort. While the clinical cohort had a smaller effect size, they did demonstrate gains from pre- to post-training in knowledge and self-efficacy. Additionally, a greater proportion of the non-clinical cohort, compared to the clinical sample, reported higher levels of awareness of risk factors for suicide after the training (93.5% vs. 82.9%). Both cohorts reported high levels of satisfaction with the training (93.3%) and 96.4% of participants found the experience valuable and the gatekeeper skill training acceptable; however, the non-clinical cohort rated the role-playing experience proportionally higher than the clinical cohort.

ASIST Program Evaluations

The Applied Suicide Intervention Skills Training (ASIST) program is a two-day workshop which aims at increasing front-line caregivers and gatekeepers’ willingness, readiness, and ability to intervene with a person at risk for suicide. The ASIST program was developed in
the early 1980s by Richard Ramsay, Bryan Tanney, William Lang, and Roger Tierney at the
University of Calgary in collaboration with the provincial and state governments of Alberta and
California and the Alberta division of the Canadian Mental Health Association
(http://www.livingworks.net/LW.php) in response to a growing concern about suicide in Alberta.
The program developers formed a partnership: Ramsay, Tanney, Tierney and Lang (RTTL) in
1983, and created LivingWorks Education Incorporated as a public service corporation in
1991. The two-day ASIST workshop is the most widely used suicide intervention training
program in the world. The program is disseminated throughout Canada, Australia, New Zealand,
United States, Norway, Northern Ireland and Scotland.

The ASIST program provides a standardized, quality controlled, learning experience. The
workshop is designed to enhance participants’ ability to provide competent life-assisting
resources for individuals at risk of suicidal behaviours. ASIST is intended as ‘suicide first-aid’ to
be applied when appropriate to individuals at imminent risk of self-harm or suicide. The training
is delivered over two consecutive days in a workshop format. Participants develop skills through
observation and supervised simulation experiences in small and large group settings. The
workshop covers five learning modules: introduction, attitudes, risk estimation, interventions
skills, and resources/networking. The structure of the workshop is fixed and participants must
attend both days of the training consecutively. The training is based on adult education principles
with less than 15 percent of the workshop using lecture format, and applies principles of
graduate learning, continuous reinforcement, and the setting of competency-based objectives.

Only 17 evaluation reports of ASIST were found. These included published and
unpublished evaluation reports (AskClyde, 2007; Bookle & Burtenshaw, 2004 (as cited in
Griesbach, Russel, Dolev & Lardner, 2008); Carney, 2005 (as cited in Griesbach et al., 2008);
Cornell, Williams, & Hague, 2006; Griesbach, Russell, Dolev, & Lardner, 2008; Guttormsen,
Høifødt, Silvola, & Burkeland, 2001; Jacobellis, 2003; MacDonald, 1999; McAuliffe & Perry,
2007; Mikhailovich, Pamphilon & Davis, 2003; Organizational Research Services [ORS], 2002,
2005; Pearce, Rickwood, & Beaton, 2003; Tierney, 1994; Turley, Pullen, Thomas, & Rolfe,
2000; Turley & Tanney, 1998; Walsh & Perry, 2000).

In general, the evaluations reported gains in participants’ suicide intervention knowledge,
skills and attitudes following training. The gains were assessed using the self-report impressions
of workshop participants in twelve evaluations and direct measures (paper tests) in four
evaluations; and simulated scenarios reported in two of the reports (Tierney, 1994; Turley et al., 2000). All evaluations reported an overall positive change in participants’ self-reported suicide intervention knowledge, skills, and positive attitudes compared to controls comprised of individuals who had not attended the workshops. The direct measures of positive attitudes, knowledge, and skills substantiated these findings and demonstrated that participants of the workshops had significant improvement post-training and in simulated scenarios when compared to pre-training, as did trained participants compared to a control group (ORS, 2002; MacDonald, 1999; Tierney, 1994; Turley et al, 2000).

In 2006, the Scottish Executive commissioned an independent evaluation of Choose Life, Scotland’s national strategy and action plan to prevent and reduce suicides (Griesbach et al., 2008). The Scottish national strategy was launched in December 2002 and as implemented in phases, with the initial phase implemented over a period of three years, from April 2003 to March 2006. The evaluation was in conformity with the growing commitment to evidence-based policy making within the Scottish government and to evidence-based practices within public health and health promotion.

Choose Life funding and local program coordination enabled 576 ASIST workshops to be delivered widely across Scotland. By September 2007, 10,477 individuals from voluntary sector projects, housing services, mental health services, primary care services, education, police, and social work had been trained — representing approximately one in five of the Scottish population (Griesbach et al., 2008).

The Choose Life program evaluation used both qualitative and quantitative methods to evaluate the breadth and depth of reactions to the ASIST program. The Kirkpatrick Model (TKM) was used as a theoretical framework for the evaluation. The program evaluation used several strategies as methodology: the national ASIST database; a national web survey of over 2000 ASIST participants; and interviews and focus groups with trainers, participants and stakeholders.

Griesbach et al. (2008) conducted an evaluation of the impact the ASIST training had in Scotland. To consider the effectiveness and impact of ASIST, the authors used The Kirkpatrick Model (TKM) as a framework for their evaluation and utilized various methods to collect their data: a national survey of ASIST participants, interviews with ASIST participants, and interviews with trainers and stakeholders. Based on self-reported measures, the authors reported
that 59% of the survey respondents reported the ASIST training challenged their attitudes about suicide. The majority of the participants (61%) who reported their attitudes were not challenged by the ASIST training indicated they found the discussion of attitudes in the ASIST useful. The proportion of participants who had reported “high” or “very high” levels of confidence, knowledge or skills in intervening with someone at risk of suicide increased substantially immediately post-training, compared to pre-training. The vast majority (76.8%) indicated their level of confidence was “high” or “very high” immediately after the training. Similarly, the majority (85.45%) of participants reported their level of knowledge was “high” or “very high” immediately after the training. Three-quarters (75.2%) of those surveyed indicated their level of intervention skills increased and was “high” or “very high” immediately after the training. Participants reported gaining knowledge of the signs that someone may be considering suicide, suicide statistics, suicide intervention model, reasons for suicide, risk factors for suicide and resources available in their community. The skills participants reported were enhanced by the training included communication skills and being able to follow through with the suicide intervention model.

The TKM evaluations also demonstrated that gains in knowledge, skills, and attitudes were maintained at follow-up (typically three to six months post training). Approximately one-fifth of the respondents had reported their knowledge and skills decreased at follow-up and a quarter of the respondents indicated their level of confidence decreased over time.

Elsewhere, in the AskClyde (2007) evaluation, an independent evaluation of the ASIST program commissioned as part of the West Dunbartonshire Choose Life Action Plan, 91% of the survey participants felt they had developed new knowledge as a result of their attendance. Only 9% reported they felt their knowledge had simply been confirmed through the workshop. The vast majority of survey participants, 77% reported they had further developed their intervention skills and 23% felt their intervention skills had been reinforced by the training. Carney (2005, as cited in Griesbach et al., 2008), in an evaluation of the ASIST in Ireland; however, found that there was a decrease in knowledge at the three months follow-up about the need to encourage distressed individuals to talk about their wish to die and the need to enquire about the distressed individual’s current life events. An annual program evaluation of the Washington State Youth Suicide Prevention Program training workshops (ORS, 2005) also reported a significant decline in the participants’ ability to correctly identify the five factors for suicide assessment. Post-
training scores decreased from 4.02 to 2.94 at the three month follow-up, and further decreased to 2.39 at the six month follow-up. An earlier program evaluation of the training workshops (ORS, 2002) found that gains in knowledge and skills over time are most sustainable among participants with pre-training experience of working with suicidal individuals. Elsewhere, McAuliffe and Perry (2007) found post training a decrease in comfort level in talking to clients about suicide.

Changes in attitudes and gains in knowledge and skills do not always correspond to a transfer and integration of the newly acquired knowledge and skills in behaviours. Ten of the studies measured the extent to which the ASIST participants had applied their acquired knowledge and skills post training. According to the self report data, nearly 50% of the participants used the model or some elements of it with at least one distressed suicidal individual within three to six months post training. Almost half of all respondents (42%) in the AskClyde (2007) evaluation reported they had put their new knowledge into practice. Among those respondents who had the opportunity to use their ASIST training, 92% indicated they had no problem in applying their ASIST training. A follow-up survey found that within six to nine months of training, the majority of respondents had applied their training in a situation with individuals who they identified as being at risk of suicide. Griesbach et al. (2008) reported 67.4% of the respondents had indicated they were much more likely to intervene with an individual at risk of suicide. The number of respondents who had indicated they had intervened with a person at risk of suicide rose from 58% pre-training to 78% post-training; however, individuals who intervened after the training were most likely to have been those who had experience of intervening prior to the training and who had also reported higher levels of confidence, knowledge and skills, prior to and after the training. McAuliffe and Perry (2007) and Cornell et al. (2006) studied the application of acquired knowledge and skills into practice into different settings.

McAuliffe and Perry (2007) evaluated the Best Practice Initiatives, of which the ASIST program was a component, implemented at Trillium Health Centre, a large two-site facility in Mississauga, Ontario. The evaluation measured: 1. the proportion of their clients the staff routinely assessed for suicide risk; 2. identification of suicidal risk among mental health patients presenting to the emergency department; and 3. admission rates of suicidal patients presenting to the emergency department. All of these measures were taken repeatedly over a four-year period.
The findings suggest that knowledge and skills were applied effectively to clinical situations. There was an increase of 14-21% in identification of suicide risk among mental health patients presenting to the emergency department, and more staff (13%) assessed a greater proportion of their patients for suicide risk. The survey also showed that the percentage of staff who knew what procedures to follow after assessing for suicide risk increased from 87 to 97%, and post-training the percentage increased to 100%; however, only 48% “strongly agreed” that they knew what to do after assessing for suicide risk. Furthermore, there was a decline (from 56 to 42%) in the number of suicidal patients’ admission rates. Study participants attributed the drop in admission rates to a clearer process of exploring reasons for dying and living and an increased focus on strengthening the patients’ protective factors and referrals to resources in the community, which enabled some admissions to be avoided.

Cornell et al. (2006) examined the transfer of knowledge into skills among primary and secondary school staff members in Virginia, USA. Over a period of two years the researchers measured: 1. the number of referrals to mental health services; 2. the number of students asked about suicide; and 3. the number of contracts not to engage in suicidal behaviour that were made with potentially suicidal students. The researchers compared the above mentioned measures pre- and post-training and with trainees versus controls. According to Cornell et al. (2006) participants made more referrals post training than they did before training. The control group; however, made more referrals than the workshop participants did. The number of students questioned about suicidal ideation did not differ pre- and post-training and compared to the control group, ASIST participants reported more incidents where they suspected a student might be suicidal and decided not to question the student (0.7 versus 6.7, respectively). Additionally, no significant differences between the number of students who contracted not to engage in suicidal behaviour were reported pre- and post-training; however, within two years post-training, the ASIST participants made more contracts with suicidal students than did the control group (Cornell et al., 2006).

Capp et al. (2001) evaluated a gatekeeper program implemented in a small Aboriginal community in Australia (n=44). The results indicated an increase in workshop participants’ knowledge of suicide, intentions to provide help, and an enhanced confidence in being able to identify those with risk of suicide. A follow-up evaluation conducted two years following the gatekeeper program (Deane et al., 2006) found that intentions to help and confidence in
identifying those at risk of suicide remained high. Over a third (37.5%) of those who participated in the workshop reported they had subsequently helped someone at risk of suicide.

Pearce et al. (2003) reported the findings of the evaluation of the Suicide Intervention Project (SIP), a multi-layered peer-based mental health program, designed as a partnership between the University of Canberra and the YWCA of Canberra. The SIP components included the ASIST program, an interactive presentation on mental health awareness, presentation about sources of support on campus, and packets of information detailing specialized support services available in the Canberra region. The program goal was to use prevention and early intervention methods as means of improving capacity to respond to suicide within the university community. The program was delivered to 56 participants of whom 42 participated in the evaluation. The pre-and post-tests indicated that the SIP had positive effect on participants’ attitudes, norms, perceived behavioural control, self-efficacy and intention towards talking to other students about feelings and mental health problems. None of these factors; however, correlated with the actual behaviour of talking to other students about feelings, which was measured two weeks after program completion.

According to Turley and Tanney (1998) qualitative data provides insight into the transfer of acquired knowledge and skills to behaviour. Many of the participants reported that the skills they acquired in the ASIST training translated into more engagement in intervention and changes in their helping activities. The changes in their behaviours were attributed to an increase in awareness of warning signs, a clearer understanding that they must intervene, an increased confidence in directly asking the distressed individual if they were suicidal, and an increased confidence in their ability to intervene.

Earlier, Tierney (1988, 1994) evaluated the Suicide Intervention Workshop (SIW), a precursor to the ASIST program. Participants’ ability to intervene with suicidal persons in simulated suicide intervention situations was evaluated with the Suicide Intervention Protocol (SIP). The SIP also measured attitude changes after a two-day workshop. Three groups of participants were used, one each for intervention abilities (n=17), knowledge (n=154), and attitudes (n=176). An equivalent control group was used exclusively for the intervention abilities segment of the study. Based on simulated suicide intervention situations, the findings indicated a significant change in intervention abilities after the workshop (p<0.001); however, workshop participants and controls did not significantly differ in their ability to recognize facilitative
intervention responses as measured by the Suicide Intervention Response Inventory (SIRI). The SIRI tested the ability of workshop participants to select facilitative responses to statements made by suicidal persons. The findings also indicated a significant enhancement in positive attitudes toward suicide interventions, general knowledge of suicide and intervention knowledge (p < 0.001). An important caveat of the SIRI is that there is often a ceiling effect which impedes the detection of significant post-training improvements. Additionally, most workshop participants retained the skills they were taught in the program six months after completion of the 2-day workshop in suicide intervention skills.

Stuart, Waalen and Haelstromm (2003) evaluated the efficacy of the Peer Gatekeeper Training (PGT) program offered to students in eight high schools in British Columbia. The PGT consisted of two-half day sessions, one week apart, with a follow-up session three months later. The skills-based training sessions consisted of the following components: a) active listening skills, b) self care and setting limits, c) crisis theory, d) suicide warning signs, e) suicide risk assessment, f) role-playing scenarios, and g) community resources (p. 324). Sixty-five adolescents participated in the study. Post training, significant gains in knowledge about suicide and skills to intervene with suicidal peers were evident. There was also an improvement in positive attitudes toward suicide intervention. The gains were maintained over a 3-month follow-up period of time.

**Methodological Limitations**

The main objective of the review of gatekeeper program evaluations was to provide a starting point for critically assessing the programs’ qualities and to judge their overall merit based on the indicators that are most relevant to this evaluation i.e., knowledge and skills acquisition, positive attitudinal changes, and the transfer of acquired skills to practice.

The extent to which conclusions can be drawn from the reviewed gatekeeper program evaluations in this chapter is limited by the relatively small numbers of existing program evaluations in general, and those published in peer-reviewed journals in particular. Concomitantly, the quality of the evidence and the heterogeneity of the methodologies (i.e, mixed-methods, repeated method design, quasi-experimental, qualitative, and quantitative), targeted groups (i.e. health care professionals, mental health professionals, education sector (high school staff, higher education professionals, community workers, youth, university
students), cultures (i.e. Australia, Scotland, Canada, USA, and Norway), participants’ socio-economic status (i.e. professionals, non-professionals), age (youth, adults), and gender (primarily female), makes it is difficult to compare the findings across evaluations. Indeed, most of the gatekeeper program evaluations reviewed in this section had numerous methodological weaknesses (e.g., small sample sizes, lack of triangulation of findings). The majority of the reviewed evaluations relied on retrospective self-reported data which can be influenced by the memory, motivation, and subjective perception of the participants. While self-reports are informative and can offer insights, they do not substitute for more direct measures.

The reported time between the training and the program evaluation varied across evaluations and in several instances was vague and unspecified (Bookle & Burtenshaw, 2004 (as cited in Griesbach et al., 2008); Gutormsen et al., 2001; Walsh & Perry, 2000). While most evaluations were conducted immediately after the training (Carney, 2005 (as cited in Griesbach et al., 2008); ORS, 2002; Pearce et al., 2003; Tierney, 1994; Turley & Tanney, 1998) others reported time lapses of one to 22 months (Cornell et al., 2006), one to twelve months (Mikhailovich et al., 2003), one to three years post-training (Walsh & Perry 2000), more ambiguously as reported by McAuliffe and Perry (2007) as on-going for four years, or not reported at all (Jacobellis, 2003).

A mixed method design was reported by five of the evaluations (Griesbach et al., 2008; Jacobellis 2003; McAuliffe & Perry, 2007; Mikhailovich et al., 2003; Walsh & Perry, 2000); however, most of these evaluations provided vague information on the quantitative measures used beyond stating “survey.” Quasi-experimental design with non-equivalent control groups and pre- and post-repeated measures was used in four of the evaluations (Cornell et al., 2006; MacDonald, 1999; Tierney, 1994; Turley et al., 2000), and while some chose to control for test effects and practice effects (Tierney, 1994) others controlled for order effects and learning effects (Turley et al., 2000). The vast majority of the program evaluations used a repeated measures design with pre and post measures (Carney, 2005 (as cited in Griesbach et al., 2008); Cross et al., 2007; Matthieu et al., 2008; McAuliffe & Perry, 2007; Mikhailovich et al., 2003; ORS, 2002, 2005; Pearce et al., 2003), while others reported only one group with post-test design (AskClyde, 2007; Griesbach et al., 2008; Walsh & Perry, 2000), and several program evaluations used only a qualitative approach such as focus groups or telephone interviews (Bookle & Burtenshaw, 2004 (as cited in Griesbach et al., 2008); Gutormsen et al., 2001). Data
collection, procedures and analyses in several evaluations was incomplete (Bookle & Burtenshaw, 2004 (as cited in Griesbach et al., 2008); Jacobellis, 2003; McAuliffe and Perry, 2007; ORS, 2002; Turley & Tanney, 1998) with several studies not providing sample sizes (Turley & Tanney, 1998) or sample characteristics (Bookle & Burtenshaw, 2004 (as cited in Griesbach et al., 2008); Carney, 2005 (as cited in Griesbach et al., 2008); Jacobellis, 2003; McAuliffe & Perry, 2007). Sample sizes in the evaluations reviewed varied and ranged between six participants (Bookle & Burtenshaw, 2004 (as cited in Griesbach et al., 2008)) to over two thousand participants (Griesbach at al., 2008). Many of the evaluations were conducted with a small group of participants (AskClyde, 2007; Bookle & Burtenshaw, 2004 (as cited in Griesbach et al., 2008); Carney, 2005 (as cited in Griesbach et al., 2008); Guttormsen et al., 2001; MacDonald, 1999; Walsh & Perry, 2000), and some evaluation reports did not provide any information about their sample sizes (Jacobellis, 2003; Turley & Tanney, 1998). None of the evaluations indicated whether participation in the training was mandatory or voluntary, nor did the majority of program evaluations describe the recruitment processes.

Very few evaluations described in detail the measures they used to assess pre- and post-training changes in their program evaluations or indicate the psychometric properties of the measures (Carney, 2005 (as cited in Griesbach et al., 2008); MacDonald, 1999; Tierney, 1994). Furthermore, several of the program evaluations had incomplete reporting of quantitative or qualitative findings (Bookle & Burtenshaw, 2004 (as cited in Griesbach et al., 2008); Guttormsen et al., 2001; Jacobellis, 2003; McAuliffe & Perry, 2007; ORS, 2002; Turley et al., 2000; Turley & Tanney, 1998; Walsh & Perry, 2000).

Although participants’ self-reported perceived competencies were high, evidence from direct measures of behavioural change was sparse and inconclusive. Four of the evaluations used either simulated scenarios (Tierney, 1994; Turley et al., 2000) or an observational assessment or a standardized role-playing dialogue (Cross et al., 2007; Matthieu et al., 2008) while all other program evaluations reviewed relied on the participants’ perception of behavioural changes and the application of skills learned to workplace situations.

In conclusion, the reviewed program evaluations had flawed methodologies; thus, they do not provide empirical evidence to support claims of program effectiveness in enhancing knowledge, skills and positive attitudes.
Summary

Despite the widespread implementation and recommendations for gatekeeper training to be incorporated into larger suicide prevention initiatives (Mann et al., 2005) the systematic evaluation of gatekeeper training programs is limited and there is a lack of consensus as to their effectiveness (Isaac et al., 2009; Mann et al., 2005).

Numerous factors affect research into the effectiveness of gatekeeper programs, among them is the low base rate of suicide which makes it difficult to attribute reductions in suicide rates to the effectiveness of a particular program (Beautrais et al., 2007). It is generally acknowledged that suicide rates are affected by a multitude of societal and individual factors (Beautrais, 1998), not just the effectiveness of suicide prevention programs. Given these complexities, it is practically impossible to attribute changes in suicide rates to a single prevention program. Additionally, there are methodological difficulties inherent in conducting suicide prevention research. Most gatekeeper training programs are implemented within a broader suicide prevention strategy; thus, isolating the unique effect of the gatekeeper program on workshop attendees’ attitudes, knowledge, and behaviour is difficult.

While some of the selected program evaluations presented in this review did not demonstrate the existence of empirical evidence to support claims of program effectiveness in enhancing knowledge, skills and positive attitudes, other gatekeeper training programs in specific institutional settings such as the military and schools have demonstrated promise (Isaac et al., 2009; Mann et al., 2005). Although, in the military the gatekeeper role is formalized, gatekeepers in other workplaces can have a critical role in increasing public awareness, encourage help seeking, improve identification of those at risk of suicide and facilitate pathways to mental health services (Beautrais, 2006). Apart from schools and the military, few suicide prevention gatekeeper programs implemented in the workplace have been evaluated (Chagnon, Houle, Marcoux, & Renaud, 2007; Cross et al., 2007; Knox et al., 2003; Matthieu et al., 2008). Further research into the effectiveness of gatekeeper suicide prevention programs implemented in the workplace is needed.
CHAPTER FOUR
METHODS

This chapter reiterates an overview of the program evaluation and outlines the research questions in depth. The rationale for the methodology used in this program evaluation is provided. This chapter describes the methods used in each phase of the evaluation, including the design, procedure, sample, data analysis and ethical considerations.

Overview of the Program Evaluation

The program evaluation consisted of a two-phase sequential mixed-method design. Phase one was a quantitative evaluation of the gatekeeper suicide prevention program: suicideAWARE and safeTALK workshops implemented at the TTC. The objective of the quantitative evaluation was to measure the impact the workshops had on participants’ factual knowledge about suicide and suicide risk factors, their attitudes toward suicide prevention, risk assessment and intervention skills.

Phase two was a mixed-method evaluation that combined a dominant qualitative component with a less-dominant quantitative component (Creswell, 2003; Teddlie & Tashakkori, 2003). The quantitative component was integrated to evaluate statistically perceived competencies and to link it to changes in behaviour in the workplace. The qualitative evaluation sought to elicit participants’ reactions and perspectives on the suicide prevention program and the impact it had on their attitudes towards suicide, and competencies to intervene with distressed suicidal patrons.

Workshops’ Objectives

The workshops were part of an education and training effort, which in turn is one part of a broader systematic effort by the TTC to deal with suicide in the Toronto subway system. The efforts involved collaborations with the Arthur Sommer Rotenberg Chair in Suicide Studies, an employee awareness campaign, and the gatekeeper suicide prevention training program.

The goal of the workshops was to provide select TTC employees with intervention competencies in order to prevent the immediate risk of injuries or death by suicide. The workshops were designed for TTC employees to develop positive attitudes toward suicide
prevention, heighten suicide awareness, increase knowledge of suicide warning signs, improve assessment of immediate risk, and enhance intervention skills and appropriate referrals of at-risk individuals.

Workshops’ Content, Activities and Format

suicideAWARE Gatekeeper Program

suicideAWARE is a general, community-oriented program designed to create a context for exploring issues in suicide prevention. The program is designed to stimulate and engage people in dialogue about suicide and to play a role in mobilizing participants to become more involved in suicide prevention activities and work towards the creation of a safer environment. The standardized program adopts an exploration format and is delivered as a half-day interactive workshop. It invites participants to consider and discuss attitudes, aided by customized video presentations that stimulate interactive involvement appropriate in small groups. The AWARE element in the workshop’s name is an acronym for Always Watch And Report Effectively. The program was modified specifically for TTC train operators and was presented as a component in their re-certification training. suicideAWARE emphasis is on creating a safe environment for exploration and discussion.

suicideAWARE has the following learning modules:
1. Magnitude of the problem
2. Attitudes and knowledge
3. Warning signs
4. Reporting
5. Self care

safeTALK Gatekeeper Program

safeTALK is a general community-oriented program modified to the needs of the TTC’s Special Constables and Mobile, Surface and Subway Supervisors. safeTALK has been standardized and presented as a full-day program that contains learning modules concerning attitudes, knowledge, intervention skills, and community resources material. The workshop utilizes a mixture of large and small group interactive formats. Safe is an acronym for Suicide Alertness For Everyone while TALK refers to Tell Ask Listen Keep safe. Following the
introduction module, the attitudes module focuses on personal experiences and personal attitudes towards suicide. The goal of the module is to reduce attitudinal barriers that may interfere with the acquisition of knowledge necessary for identification of suicidal behaviour and intervention skills. The attitudinal module encourages participants to examine their own attitudes and the impact they may have on responding to suicidal patrons. The knowledge module is presented after the attitudes module and is intended to develop the capacity to inform others about the importance of intervention and to enhance knowledge about warning signs of suicide. The module also promotes ability to communicate this information to others, and identify behaviours that require intervention. The fourth module focuses on intervention skills acquisition and assists participants in enhancing their intervention skills. Each phase of the intervention involves specific tasks: 1) the exploration phase involves identification (invitation and Tell) and engagement (Ask), followed by 2) the understanding phase (Listen) which involves inquiry and assessment tasks, and lastly, 3) the action phase (Keep safe) which includes implementation tasks and referral to appropriate community resources. The full-day interactive workshop includes discussions, didactic teaching, video presentations, and role-playing.

Program Evaluation Questions

As previously indicated in Chapter One, a number of research questions were outlined for this evaluation, which can be answered through the overall mixed methods design, encompassing phases one and two.

1. Are the suicide prevention program objectives achieved?
   a) Do participants’ positive attitudes increase as a result of the workshop?
   b) Do participants acquire increased knowledge?
   c) Are participants’ interventions skills enhanced by the training, and are participants able to apply the intervention skills? Have the skills learned been applied to interventions?
   d) Are the enhanced attitudes, acquired knowledge, and intervention skills retained over three months?

2. Are there immediate training effects on perceived intervention capabilities, comfort, and confidence?
TTC Program Evaluation Methodological Approach

The Kirkpatrick Model (TKM) provided the analytical framework guiding the two-phase, sequential mixed-model approach to the gatekeeper program evaluation. The program evaluation consisted of two main phases. The two phases of the evaluation followed each other sequentially with the qualitative phase of the evaluation following the quantitative phase. As illustrated in Table 1, the two phases of the evaluation measured different aspects of the training experience which corresponded to different levels in the analytical framework: reactions (Level I of TKM), learning (Level II of TKM), and behaviour (Level III of TKM). Phase one was a quantitative evaluation which examined the degree TTC employees who attended the suicide prevention workshops were sensitized to their role as gatekeepers i.e., increased in factual knowledge about suicide and suicide risk factors, enhanced positive attitudes toward suicide and suicide intervention, and improved risk assessment and intervention skills.

In the second phase, the analytical framework guided the face-to-face interviews designed to elicit employees’ reactions to the gatekeeper suicide prevention program, and explored the impact the program had on their competencies to intervene with distressed TTC patrons.

Rationale for the Mixed-method Approach

Johnson and Onwuegbuzie (2004, p. 20) distinguish between mixed-model and mixed-method designs with the former design referring to mixing quantitative and qualitative approaches within or across the stages of the research process, and the latter referring to the inclusion of a quantitative phase and a qualitative phase in an overall research study. The present study used a sequential mixed-methods design with a quantitative phase followed by a qualitative phase.

The underlying logic of mixing methods is that neither quantitative nor qualitative methods are sufficient in themselves to capture the intricacies of a program; moreover, there are strengths and weaknesses to any single data collection strategy. Using more than one data collection approach permits the researcher to combine strengths and correct some of the deficiencies of any one source of data (Greene, Caracelli, & Graham, 1989; Patton, 1987; Rossman & Wilson, 1985). When used in combination, both quantitative and qualitative data...
yield a more complete analysis and add complexity to the evaluation design and use the advantages of both the quantitative and qualitative approaches (Creswell, 1994, 2003; Gogolin & Swartz, 1992; Greene, Caracelli, & Graham, 1989; Grinnell, 1997; Johnson & Onwuegbuzie, 2004; Kushman, 1992; Madey, 1982; Patton, 1987; Rossman & Wilson, 1985; Teddlie & Tashakkori, 2003). Although the two approaches are seen as complimentary, they represent distinct universes (Rossman & Wilson, 1985).

No single method ever adequately solves the problem of rival causal factors. Because each factor reveals different aspects of empirical reality, multiple methods of observation must be employed (Denzin, 1978, p. 28).

Table 1: Summary of the Kirkpatrick Model (TKM)

<table>
<thead>
<tr>
<th>Level</th>
<th>Measure</th>
<th>Key Question</th>
<th>Methodologies or Indicators</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reaction</td>
<td>Satisfaction</td>
<td>What was the participants' reaction to the program?</td>
<td>Interviews, questionnaires, and general participant comments throughout the program</td>
<td>Not a Randomized Control Trial. Can't measure what is learned, guarantee behavioural change, nor will it ensure quality results from learning</td>
</tr>
<tr>
<td>Learning</td>
<td>Knowledge</td>
<td>What did the participants learn?</td>
<td>Pre-post testing, interviews and surveys</td>
<td>Will not measure if they liked the program, if they will behave differently, and if expected results will be achieved</td>
</tr>
<tr>
<td>Behaviour</td>
<td>Transfer of learning &amp; Achievement of performance objectives</td>
<td>Did the participants' learning affect their behaviour?</td>
<td>Interviews</td>
<td>Cannot determine if participants liked the training, and if the behaviours accomplished results</td>
</tr>
<tr>
<td>Results</td>
<td>Transfer or impact on society</td>
<td>Did participants' behaviour changes affect the suicide rates in the subway system?</td>
<td>Indicators include increases in MHA apprehensions and reduced suicide incidents in the subway system</td>
<td>Cannot determine if participants liked the training, understood it, or if it affected their preferred behaviours</td>
</tr>
</tbody>
</table>
The use of both quantitative and qualitative methods is analogous to “zooming in and zooming out with a lens” and allows the researcher to focus on different aspects of the same phenomenon (Willems & Raush, 1969, pp. 82-83). While the qualitative evaluation provides depth to some of the causes behind changes and offers additional explanations, the quantitative methods provide a yardstick for measuring change. Madey (1982, p. 235) suggests that “when selecting methods and designing a study, evaluators should be receptive to the enormous potential that integrating methods offers to improving studies.”

There are extensive arguments on how mixed methods allow for improvement of the accuracy of conclusions by relying on data from more than one method. Greene, Caracelli and Graham (1989), for example, proposed five major rationales for the integration of mixed-model approaches in program evaluations:

- **Triangulation** - it allows for the examination of data collected through more than one method to seek convergence in the findings.
- **Complementarity** - it permits elaboration, enhancement, clarification, and illustration of results from one method with results from the other method. It provides richness in detail expanding understanding of the participants’ reactions to the program and perceived relevance to their work. It explores interconnected and/or distinct aspects of the data.
- **Initiation** - it facilitates the discovery of inconsistencies and contradictions in the data that possibly lead to a substantial alteration and re-framing of the research question.
- **Development** - it allows for augmentation by using the findings from one method to help inform and/or develop the other method which enhances the validity of the constructs by capitalizing on the inherent strengths of the method.
- **Expansion** - it increases the breadth and depth of the results and interpretation by analyzing them from different perspectives.

There were many advantages to adopting a mixed methodology for this program evaluation that were particularly relevant to the overall aims of the evaluation. It was argued that the objectives outlined earlier were best answered using a mixed methods approach. Mono-method approach was enhanced, elaborated on and weakness and/or limitations were decreased by the addition and the juxtaposing of an alternative method. The qualitative methods augmented quantitative components and provided richness and completeness to the evaluation by integrating participants’ perspectives and impressions of the program.
Phase two used mixed-method design that combined a dominant qualitative component with a less-dominant quantitative component. The “dominant-less dominant” mixed-method research design refers to research in which “one paradigm and its methods predominate, with a smaller component of the overall study being drawn from an alternative design” (Tashakkori & Teddlie, 1998, p. 44). In phase two the quantitative and qualitative data were collected concurrently. The smaller quantitative component was integrated in the qualitative interview to evaluate statistically perceived competencies and link them to changes in behaviours in the workplace.

The Logic Model illustrated in Appendix C, is the means of describing the program theoretically from a program evaluation perspective. The Logic Model was part of the planning of the gatekeeper suicide prevention initiative at the TTC and it identified change prospectively. It helped plan the program with the focus on the ultimate outcome i.e., the reduction of suicide incidents in the Toronto subway system. The Logic Model depicts the causal processes through which change occurred as a result of the program’s strategies and actions. As part of the evaluative process, the Logic Model identifies change retrospectively (Shapiro, 2005) and provided the intermediate or shorter-term indicators, in terms of outcomes to measures the program’s effectiveness.

**Phase 1: TKM Level II - Learning**

**Quantitative Evaluation**

**Design**

To test the program evaluation hypotheses 1, 2, and 3 the quantitative evaluation used a repeated measures design. The repeated measures design consisted of self-reported pre-training questionnaires and similar self-reported post-training questionnaires. The repeated measures design permitted an evaluation of the workshop’s effects on participants’ knowledge acquisition, attitudinal changes, and intervention skills enhancement. Long-term knowledge retention was tested three months after workshop attendance.
Sample

All TTC employees attending the safeTALK and suicideAWARE workshops were invited to voluntarily participate in phase one of the evaluation by completing the pre and post questionnaires (Appendices D, E, F, and G).

The safeTALK and suicideAWARE workshops evaluated in this program evaluation were delivered to TTC employees from September 28th, 2005 to September, 30th, 2006. The workshops were presented under the auspices of the TTC Training Department and delivered by trainers from the Trillium Health Centre. The 13 safeTALK workshops were delivered to 176 TTC employees in the last quarter of 2005, from September 28th to Dec. 17th. The workshops were offered to transportation supervisors (Mobile, Route, Surface, Subway), Chief Supervisors, TTC Special Constables, instructors, safety personnel, and occupational health personnel. As indicated in Table 2, of the 176 safeTALK participants, 136 (77%) completed the pre-training questionnaires, and 137 (78%) completed the post-training questionnaires. Follow-up questionnaires were mailed to 136 workshop participants who completed the pre-training questionnaires: 107 follow-up questionnaires were completed and returned. The effective response rate; therefore, was 78.7%.

As shown in Table 4, participants in the safeTALK workshop were from different occupational groups within TTC: 43% were Supervisors (18% Subway Supervisors, 23% Mobile and Surface Supervisors, 2% Chief Supervisors), 16% were Special Constables, and 41% of the participants were from other TTC departments (Training, Safety and Occupational Hygiene, Human Resources). The majority of safeTALK participants were males (77%), 35 years of age and older (87%), and with 10 or more years of employment at the TTC (62%).

The suicideAWARE workshops were presented as part of the re-certification training offered by the TTC Training Department to all Train Operators. The re-certification training is offered bi-annually and is compulsory for all Train Operators. Thirty-five suicideAWARE workshops delivered from January 1st, 2006 to Sept 30th, 2006. As illustrated in Table 2, of the 184 Train Operators who attended the workshops, 171 (93%) completed the pre-training questionnaires; and 172 (93.5%) completed the post-training questionnaires. Follow-up questionnaires were mailed to the 171 participants who completed the pre-training questionnaires; 67 questionnaires were completed and returned. The effective response rate; therefore, was 38.9%. Table 5
illustrates Train Operators’ participation in the program evaluation. Of the 171 Train Operators who had completed the pre-training evaluation questionnaire, the majority were males (92%), 35 years of age or older (82%) with 10 or more years of employment at the TTC (73%).

Overall, as shown in Table 2, the vast majority of those who attended the suicide prevention workshops completed the pre and post questionnaires (85.3%; 307/360). Follow-up questionnaires were mailed to 307 participants and 174 questionnaires were completed and returned. The effective response rate; therefore, was 56.7%. As indicated in Table 3, the majority of the workshops participants were male (85.6%), over the age of 35 (84.3%), and over half of the sample participants (68.2%) were employed by the TTC for over 10 years, with the majority of the study participants (55.4%) employed as Train Operators. The study sample is representative of the overall TTC employees’ demographics in term of gender distribution, age, and years of employment (V. Cosentino, personal communication, July 27, 2010),

Table 2: Summary of Participation in the Program Evaluation

<table>
<thead>
<tr>
<th>Participation</th>
<th>safeTALK</th>
<th>suicideAWARE</th>
<th>All Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Percent</td>
<td>N</td>
</tr>
<tr>
<td>Number of workshop participants</td>
<td>176</td>
<td>100%</td>
<td>184</td>
</tr>
<tr>
<td>Completed pre-training questionnaires</td>
<td>136/176</td>
<td>77.30%</td>
<td>171/184</td>
</tr>
<tr>
<td>Completed post-training Questionnaires</td>
<td>137/176</td>
<td>77.80%</td>
<td>172/184</td>
</tr>
<tr>
<td>Completed follow-up questionnaire</td>
<td>107/137</td>
<td>78.10%</td>
<td>67/172</td>
</tr>
<tr>
<td>Linked questionnaires</td>
<td>63/137</td>
<td>46.00%</td>
<td>65/169</td>
</tr>
</tbody>
</table>
Table 3: Characteristics of 305 Workshop Participants

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>N</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25-34</td>
<td>48</td>
<td>15.30%</td>
</tr>
<tr>
<td>35-49</td>
<td>161</td>
<td>52.80%</td>
</tr>
<tr>
<td>50-65</td>
<td>96</td>
<td>31.50%</td>
</tr>
<tr>
<td><strong>Gender</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>261</td>
<td>85.60%</td>
</tr>
<tr>
<td>Female</td>
<td>44</td>
<td>14.40%</td>
</tr>
<tr>
<td><strong>Position</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Special Constables</td>
<td>22</td>
<td>7.20%</td>
</tr>
<tr>
<td>Transportation Supervisors</td>
<td>58</td>
<td>19.00%</td>
</tr>
<tr>
<td>Train Operators</td>
<td>169</td>
<td>55.40%</td>
</tr>
<tr>
<td>Other</td>
<td>56</td>
<td>18.40%</td>
</tr>
<tr>
<td><strong>Years of Employment at TTC</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 10 years</td>
<td>97</td>
<td>31.80%</td>
</tr>
<tr>
<td>&gt; 10 years</td>
<td>208</td>
<td>68.20%</td>
</tr>
</tbody>
</table>

**Linked Data**

When participants indicated their TTC employee identification number on the evaluation questionnaires, their pre, post and follow-up questionnaires were linked as a complete data set. As shown in Table 4, of the 136 participants who attended the *safeTALK* workshop, 63 indicated their TTC employee identification number on all three questionnaires; thus, 46.3% of pre, post and follow-up questionnaires were linked. As illustrated on Table 4, no significant differences between sample characteristics of the linked and unlinked data were observed in participants of the *safeTALK* workshop with the exception of occupation (0.032) with a greater number of Transportation Supervisors with unlinked data (22 vs. 36).
Of the 171 Train Operators who have completed the evaluation questionnaires only 65 participants indicated their TTC employee identification numbers on the evaluation questionnaires; thus only 38% of the pre, post and follow-up questionnaires were linked.

Table 4: Sample Characteristics of safeTALK Workshop

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Linked Data (n=63)</th>
<th>Unlinked Data (n=73)</th>
<th>All Participants (n=136)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Percent</td>
<td>N</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25-34</td>
<td>12</td>
<td>19.00%</td>
<td>6</td>
</tr>
<tr>
<td>35-49</td>
<td>35</td>
<td>56.00%</td>
<td>47</td>
</tr>
<tr>
<td>50-65</td>
<td>16</td>
<td>25.00%</td>
<td>20</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>49</td>
<td>78.00%</td>
<td>56</td>
</tr>
<tr>
<td>Female</td>
<td>14</td>
<td>22.00%</td>
<td>17</td>
</tr>
<tr>
<td>Position</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Special Constables</td>
<td>10</td>
<td>16.00%</td>
<td>12</td>
</tr>
<tr>
<td>Transportation</td>
<td>22</td>
<td>35.00%</td>
<td>36*</td>
</tr>
<tr>
<td>Supervisors</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Others</td>
<td>31</td>
<td>49.00%</td>
<td>25</td>
</tr>
<tr>
<td>Train Operators</td>
<td>0</td>
<td>0%</td>
<td>0</td>
</tr>
<tr>
<td>Years of Employment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>at TTC</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 10 years</td>
<td>27</td>
<td>43.00%</td>
<td>25</td>
</tr>
<tr>
<td>&gt;10 years</td>
<td>36</td>
<td>57.00%</td>
<td>48</td>
</tr>
</tbody>
</table>

* Comparison of Linked vs. unlinked for occupational group Transportation Supervisors’ characteristics p= 0.035

Sample characteristics are available for 305 participants; two participants chose not to complete the demographics page
Table 5: Sample Characteristic of suicideAWARE Workshop participants

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Linked Data (n=65)</th>
<th>Unlinked Data (n=106)</th>
<th>All Participants (n=171)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Percent</td>
<td>N</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25-34</td>
<td>12</td>
<td>18.00%</td>
<td>19</td>
</tr>
<tr>
<td>35-49</td>
<td>31</td>
<td>48.00%</td>
<td>48</td>
</tr>
<tr>
<td>50-65</td>
<td>22</td>
<td>34.00%</td>
<td>37</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>59</td>
<td>91.00%</td>
<td>97</td>
</tr>
<tr>
<td>Female</td>
<td>6</td>
<td>9.00%</td>
<td>7</td>
</tr>
<tr>
<td>Position</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Special Constables</td>
<td>0</td>
<td>0%</td>
<td>0</td>
</tr>
<tr>
<td>Transportation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Supervisors</td>
<td>0</td>
<td>0%</td>
<td>0</td>
</tr>
<tr>
<td>Others</td>
<td>0</td>
<td>0%</td>
<td>0</td>
</tr>
<tr>
<td>Train Operators</td>
<td>65</td>
<td>100.00%</td>
<td>106</td>
</tr>
<tr>
<td>Years of Employment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>at TTC</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 10 years</td>
<td>21</td>
<td>32.00%</td>
<td>24</td>
</tr>
<tr>
<td>&gt;10 years</td>
<td>44</td>
<td>68.00%</td>
<td>80</td>
</tr>
</tbody>
</table>

Procedure

The pre-training evaluation questionnaires were distributed by Trillium Health Centre trainers to all workshop attendees as they arrived to attend the workshop. Prior to each training workshop, the Superintendent of Training addressed the workshop participants and explained the evaluation objectives. Completion of the evaluation questionnaires took approximately 20 minutes. The evaluation questionnaires were stapled together (Appendices D and E). Materials
were handed out until 10 minutes past the posted starting time for the workshop. Those who arrived after this time were not invited to participate as their involvement would cause undue delay in beginning the workshop. Five percent (18/360) of the workshop attendees did not participate in the study because of their delay getting to the training session. The post-training evaluation questionnaires were handed out by the trainers at the close of each training workshop (Appendices F and G). The completed questionnaires were returned to the workshop trainers and all completed questionnaires were returned in a sealed envelope to the TTC Training Department. Participation in the evaluation program was voluntary and was clearly stated on the front page of the questionnaire packet. The issue of confidentiality was also addressed. The first page of the packet included a short demographic information page and at the top of the page space was provided for the participant to write down their TTC employee identification number.

The long-term effects of suicideAWARE and safeTALK training programs on participants’ knowledge, attitudes, and suicide intervention abilities were evaluated three months after attendance in the programs. All participants received packets that included a letter to the employee, a questionnaire and a stamped pre-addressed envelope (Appendices H and I). The letter addressed to the employee stated the goals of the evaluation, addressed confidentiality and anonymity issues, and re-stated the voluntary nature of participation. The follow-up evaluation packets were sent to workshop participants through TTC internal mail. Non-responding participants were mailed a second request six weeks later.

**Measures**

Table 6 shows a summary of all measures used in phase one of the program evaluation. Each measure was designed to evaluate a key component of the change observed in workshop participants. Instruments measuring knowledge acquisition and retention and attitudinal change were adapted from instruments previously used to evaluate suicide prevention programs. The measures listed in Table 6 were viewed as most appropriate and judicious given the time constraints. The questionnaires conveyed both the instructions and the questions to participants and supplied areas for them to complete their answers. Table 7 provides a summary of target occupational groups and measures administered at the attended workshop.
Table 6: List of Outcome Measures Used in the Evaluation

<table>
<thead>
<tr>
<th>Measure</th>
<th>What it measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Suicide Opinion Questionnaire (SOQ)-Modified (Domino, Moore, Westlake, &amp; Gibson, 1982)</td>
<td>Attitudes toward suicide and the suicidal person</td>
</tr>
<tr>
<td>Suicide-risk Procedural Questionnaire</td>
<td>Participants’ knowledge of procedural actions that need to be initiated in response to potential at-risk behaviour</td>
</tr>
<tr>
<td>Intervention Knowledge Test (IKT)-Modified (Tierney, 1988)</td>
<td>Intervention knowledge</td>
</tr>
<tr>
<td>Suicide Intervention Response inventory (SIRI)-Modified (Neimeyer &amp; Bonnelle, 1997)</td>
<td>Ability to recognize facilitative responses in suicide intervention situations</td>
</tr>
</tbody>
</table>

Table 7: Summary of Targeted Occupational Groups, Workshop, and Evaluation Measures

<table>
<thead>
<tr>
<th>Employee Group</th>
<th>Workshop</th>
<th>Workshop Duration</th>
<th>Measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>TTC Special Constables</td>
<td>safeTALK</td>
<td>Full-day</td>
<td>SOQ-Modified SPQ IKT-Modified SIRI-Modified Interviews</td>
</tr>
<tr>
<td>Transportation Supervisors</td>
<td>safeTALK</td>
<td>Full-day</td>
<td>SOQ-Modified SPQ IKT-Modified SIRI-Modified Interviews</td>
</tr>
<tr>
<td>Others</td>
<td>safeTALK</td>
<td>Full-day</td>
<td>SOQ-Modified SPQ IKT-Modified SIRI-Modified Interviews</td>
</tr>
<tr>
<td>Train Operators</td>
<td>suicideAWARE</td>
<td>Half-day</td>
<td>SOQ-Modified SPQ</td>
</tr>
</tbody>
</table>
Knowledge (Hypothesis 1)

Increase in knowledge was assessed using the Intervention Knowledge Test (IKT) (Tierney, 1988) and a questionnaire designed by this researcher for the purpose of this evaluation.

Intervention Knowledge Test (IKT)

The IKT was designed and developed to measure knowledge components directly related to the ASIST workshop. Knowledge items generally focused on material related to interventions. The IKT was designed to measure salient aspects or learning points in three major modules of the ASIST workshop: attitudes, knowledge, and intervention skills. In its original format, the IKT consists of 20 multiple choice questions and two case study questions requiring assessment of suicide risk. The IKT has a moderate level of internal consistency with a split-half reliability of $r = .59$ and test-retest reliability $r = .67$ indicative of moderate stability (Tierney, 1994;1988).

For the purpose of this evaluation, the questions were modified to reflect TTC procedural requirements (i.e., question number 14: “It is important to move through a suicide intervention process while (circle all that apply); a) keeping self and others safe by securing the platform; b) ensuring trains return to service quickly; c) communicating with transit control; d) all of the above; e) none of the above”) and safeTALK workshop material (i.e., question number 13: “Key tasks in safeTALK suicide Intervention Model are: a) ask, listen, keeping safe; b) keeping safe, ask, listen; c) listen, keeping safe, ask; d) none of the above”). The modified version consisted of 17 multiple choice questions. The case study questions were eliminated as were questions directly related to the ASIST workshop material. Scores on the Intervention Knowledge Questionnaire-Modified (IKT-M) were computed by combining the correct responses with a maximum possible total score of 17. A copy of the modified IKT is provided in Appendix D & E.

Suicide-risk Procedural Questionnaire (SPQ)

Participant’s knowledge of behaviours that may be indicative of suicide intent was measured with the Suicide-risk Procedures Questionnaire (SPQ). The questionnaire was developed after extensive consultations with clinicians and TTC employees who have had
experience with suicidal patrons. The questionnaire consisted of four multiple choice questions describing behaviours that a distressed TTC patron may exhibit and may be indicative of suicidal intent. In addition to the identification of at-risk behaviour, the questions also referred to procedural actions that need to be initiated by the TTC employee upon observing such behaviour. Scores on the SPQ were computed by combining the correct responses with the maximum possible total score of four. A copy of the SPQ is provided in Appendix D & E.

Attitudes (Hypothesis 2)
Suicide Opinion Questionnaire (SOQ)

Attitudes were assessed using a modified version of the Suicide Opinion Questionnaire (SOQ) (Domino et al., 1982). The SOQ has been used extensively by researchers (Domino, 2005). The original version of the SOQ consists of 100 items that cover a wide range of attitudes towards suicide. For this evaluation several revisions were made to the SOQ in order to streamline the measure and to make it relevant for use with TTC employees. The revised SOQ-R is presented in the Appendix D & E. The primary revision to the SOQ involved streamlining the SOQ and selecting only 20 items from the standard 100 items. Streamlining the amount of time required by each participant to complete the comprehensive set of research measures, twice within the time allotted for training, was the main consideration taken into account when revising the SOQ. Identification of items that had relevance to the TTC employees and specifically to the material covered in the workshops and the printed materials distributed to TTC employees was the second consideration. The revision to the SOQ involved identifying items that pertain to factual information, misconceptions, or myths about suicide and excluding items that addressed an individual right-to-die, items that dealt with the relationship between suicidal behaviour and religion, items that referred to suicidal behaviour as inwardly or outwardly directed aggression, or items that referred to suicidal behaviour as a serious moral transgression. The SOQ questions were carefully read and crossed referenced with the printed material distributed to TTC employees with their salary statement and published in the TTC internal newsletter, the Coupler. The SOQ items were also crossed referenced with the material cover in the workshops. Additionally, the modified SOQ used in this evaluation excluded items that were not found to contribute meaningfully to the factor scales reported in the research (Domino et al., 1988-89). A lack of consensus regarding the scoring and interpretation of the
SOQ evident in the literature (Rogers & DeShon, 1992) particularly in light of the lack of internal consistency reliability estimates and evidence of validity, contributed to the decision to include selected items only. Each item in the modified SOQ was rated on a five-point Likert scale from “Strongly agree” to “Strongly disagree.”

**Intervention Skills (Hypothesis 3)**

**Suicide Intervention Response Inventory (SIRI)**

Enhancement of *safeTALK* participants’ suicide intervention ability was measured with a modified Suicide Intervention Response Inventory (SIRI) (Neimeyer & Bonnelle, 1997). Neimeyer and MacInnes (1981) described the SIRI as “designed to measure a care-giver’s competence in discriminating between facilitative and non-facilitative responses to the suicidal client” (p.176). The original instrument consists of 25 hypothetical client statements and provides two possible helper responses for each of these client statements. One response is considered facilitative for suicide prevention whereas the other one is neutral or deleterious to effective intervention. The scoring uses a 7-point Likert scale (-3 to +3) for the appropriateness of each response. The +3 indicated a highly appropriate response and -3 a highly inappropriate response. The SIRI was used by Neimeyer and Bonnelle (1997) to assess counselling skills of professional and paraprofessional counsellors including suicide intervention volunteers. The internal consistency was reported to be high with alpha coefficients of .90 and .93 at Time 1 and Time 2 respectively and test-retest reliability $r = .92$, $p < .001$ (Neimeyer & Bonnelle, 1997).

The SIRI was modified and streamlined for this evaluation. The original SIRI statements were modified to reflect situations that may occur in the context of the TTC work environment. The modified version included only 6 hypothetical client statements. The six hypothetical patrons statements were followed by two replies (Helper A and Helper B), one of which demonstrates relatively more skillful management of the suicidal patron. Participants were asked to identify which of the helpers’ statements were facilitative intervention responses and which statements were deleterious intervention responses. Facilitative responses were to be assigned a positive value (+1 to +3) and deleterious responses were to be assigned a negative value (-1 to -3). To confirm which of the helpers statements were facilitative and which were deleterious the SIRI-R questionnaire was distributed among several clinicians in the Mental Health Services at St. Michael’s Hospital. The clinicians were asked to identify among the helpers’ twelve replies
those who indicated a more skillful management of the suicidal patron. There was consensus among the clinicians regarding the facilitative and deleterious helpers’ replies: however, there was no consensus among the clinicians the degree the helper response was facilitative (+1, +2 or +3) or conversely, deleterious (-1, -2 or, -3). Thus the scoring of the participants’ responses was based on the replies identified by the clinicians as either facilitative or deleterious. The total score on the SIRI-R represents the number of correct responses the participant identified. Hence, a higher score reflects more awareness of appropriate suicide intervention skills.

A copy of the modified Suicide Intervention Response Inventory is provided in Appendix E.

**Data Analysis**

This section outlines the procedures and rationale for the quantitative data analysis conducted in phase one of the evaluation.

Initially, pre, post and follow-up data collected from participants were classified into two groups according to the workshop attended i.e., safeTALK and suicideAWARE and analyzed separately with comparison between linked and unlinked data; however, after extensive consultation with a statistician, to corroborate the data analyses approach and ensure optimal utilization of the data, it was decided that combining the data of all participants into one data set would increase the statistical power of the analyses.

All pre, post, and follow-up responses were analyzed using descriptive statistics (frequencies, mean, median, and standard deviation). In all of the statistical analyses performed, the study outcome measures i.e., knowledge, attitudes and skills were considered dependent variables while “time” (pre, post, and follow-up) was considered the independent variable.

The generalized estimating equations (GEE) were used for analyzing repeated categorical data. GEE estimates were obtained by using the exchangeable working correlation. The GEE technique with alternating logistic regressions was used to account for the correlation of observations within each subject. The generalized estimating equation (GEE) approach (Liang & Zeger, 1986) is an extension of generalized linear models that provides a semi-parametric approach to longitudinal data analysis with univariate outcomes for which the quasi-likelihood formulation is sensible, for variables such as normal, Poisson, binomial, and gamma response. GEE is the estimating equation that accommodates the correlation structure of the repeated measurements. This approach encompasses a broad range of data situations, including missing
observations, continuous explanatory variables, and time-dependent explanatory variables. The exchangeable correlation structure, which is similar to compound symmetry structure, assumes that the correlations are equal across time points.

As all of the scores were continuously distributed responses, linear mixed-effects models (LMEs) were used to describe the immediate and long-term effects of the training on attitudes, knowledge acquisition and skills enhancement. The LME approach provides a flexible yet parsimonious way of modeling the association among repeated measurements. Linear mixed-effects models (LMEs) were chosen because they are based on less restrictive assumptions than general linear models. Among the assumptions that are relaxed in LMEs are those related to the correlations between observations. LMEs allow flexible modeling of the covariance structure of the data and thus can adequately model data in which the observations are not independent. LMEs are also less restrictive in that they can be applied and deal effectively with missing data through likelihood based estimation. The simple fixed-effects structure includes a three-level factor representing time (independent variable). Prior to fitting these models, preliminary analyses were conducted to verify the appropriateness of the approach. As the design comprised only three time points, there was limited choice of error structures for the residuals. A comparison of goodness of fit using Akaike’s Information Criteria and Schwartz’s Bayesian Criteria of two structures (compound symmetry and unstructured error structure variance-covariance matrix) was conducted to select the best correlation structure.

While the unstructured covariance structure is parameterized directly in terms of variance and covariance where the observations for each pair of times have their own unique correlation, compound symmetry (CS) is the covariance structure with the simplest correlation model; it assumes the correlation is constant regardless of the distance between the time points. The CS structure is appealing because it is relatively simple, requiring the estimation of two variance components. The compound symmetry structure (CS) on the residuals allows for within-subject correlations across time. CS structure has constant variances across time and constant covariance between measurements.

To assess the impact of occupation and years of employment on the outcome measure, secondary repeated measure models were performed adding a term for: 1. years of employment at the TTC by time, and 2. Occupation by time (independent variables). These models removed the assumption of equal effects over time. Contrast tests permitting post-hoc pair-wise
comparisons between time and group means were performed primarily using Bonferroni and Sidak approaches. The Hochberg correction for multiple comparisons (Hochberg, 1988) was also used and no differences in p value were detected.

For each outcome variable, the F statistics for the overall time effect is reported, and when the effect is significant, the results of contrast testing for change between pre to post, post to follow-up and pre to follow-up are also provided.

To contextualize the effects of the training on knowledge, skills and attitudes acquisition, the magnitude of training effects was computed with Cohen’s d (Cohen, 1977;1965) for repeated measures. The formula used for the computation of the effect size reported in Chapter Five can be found in Appendix K. All the Cohen’s effect size calculations for repeated measures were conducted by the calculator found at the following website: http://wilderdom.com/courses/calculators/Cohensdrepeatedmeasures.xls. To facilitate the interpretation of the effect size, 95% confidence intervals were also computed to examine whether the average effect size encompasses zero. They are provided where applicable. Cohen (1977;1965) provides a rough guideline for the interpretation and evaluation of d statistic: with d = 0.2 (small effect), d = 0.5 (medium effect), and d = 0.8 (large effect).

The effect size (ES) reflects the magnitude of the difference between the pre and post training and pre and 3 months follow-up which is expressed in standard deviation units. Effect size can be thought of as “the degree of departure from the null hypothesis” (Cohen, 1977, p. 20). The null hypothesis in this evaluation would be that the training did not produce any gains in knowledge, positive attitudes or skills.

Numerous methodologists have advocated that researchers report measures of effect size (ES) that go beyond the significance of testing models rather than simply rely on p-values as the sole criterion of an experiment’s success (Cohen, 1965, American Psychological Association, 2001). There are many ways to compute effect size; none of which have been chosen as a gold standard (Wolf, 1986). Another method to examine the effect of the training is to compute the magnitude of the ratio between the estimated marginal means and standard error, where the estimated marginal means refers to unweighted means.

The percentage of change in mean scores of post-training scores, 3-month follow-up and pre-training scores were also calculated and provided in Chapter Five. A percentage change is a way to express the magnitude of change in a variable and it represents the relative change
between the old value and the new one. The percentage of change was calculated by subtracting
the pre-training mean score from the post or follow-up score and dividing by the pre-training
score and multiplying the result by 100.

Changes to the coding and scoring of the modified Suicide Opinion Questionnaire (SOQ-M)
and the modified Suicide Intervention Response Inventory (SIRI-M) were necessary.
Initially, the responses to individual SOQ-R items were coded according to numerical
equivalents (i.e. “strongly agree” corresponded to 1 and “strongly disagree” to 5), and two scales
were constructed: Factual Knowledge scale which consisted of eleven items and a
Favourable/Unfavourable Attitude scale which consisted of six items. Also, scores on six items
of the SOQ-R were reversed (i.e. “strongly agree” corresponded to 5) with four of the reversed
score items included in the Favourable/Unfavourable Attitudes scale. The Factual scale and
favourable attitudes scale scores were computed by combining the scores for the selected SOQ-R
items that comprised each scale. Factor analysis using correlation model, iterative criteria with
VARIMAX rotation was conducted. The results failed to support the 2-factor model. The lack of
a stable latent structure of the modified SOQ could have been caused by the streamlining of the
number of items in the modified questionnaire from 100 to 20. Further support for the lack of
sound psychometric properties of the SOQ was recently provided by Anderson, Lester and
structural validity of three competing models used extensively in research settings (the 15-factor
model, the 8-clinical scale model, and selected factors from the 5-factor model) in an attempt to
determine whether the SOQ had a psychometrically sound interpretive structure. The
confirmatory factor analyses provided no support for any of the aforementioned models. The
exploratory factor analysis of the data resulted in a weak 2-factor structure (Factual knowledge
and Acceptability of Suicide), which accounted for only 15.3 percent of the variance; however, a
subsequent study failed to support the 2-factor model.

As the confirmatory factor analyses failed to support the 2-factor model
(Favourable/Unfavourable Attitudes sub-scale and Factual Knowledge sub-scale) the responses
to individual SOQ-M items were coded as positive or negative (1 and 0, respectively) and a
participant’s total score consisted of the number of positive attitudes endorsed. The maximum
possible score for the SOQ-M was 20.
Participants’ responses to SOQ-R items were further evaluated through the calculation of “a percentage of endorsement statistic” to yield data concerning levels of agreement, disagreement and undecided responses for individual SOQ-R items. Responses of “strongly agree” and “agree” were combined and scored as “agree.” The “strongly disagree” and “disagree” responses were scored as “disagree.” Lastly, the “undecided” responses were recorded as indicated.

The Suicide intervention Response Inventory consisted of six hypothetical patron’s statements that may be indicative of distress and suicide risk, and provides two possible helper responses for each patron’s statements. Facilitative responses were to be assigned a positive value of +1 to +3 whereas deleterious responses were to be assigned a negative value of -1 to -3. Due to the lack of consensus among clinicians whether a response merited a +1, +2 or +3 or conversely -1, -2, or -3 the scoring format was changed. The new scoring format was based on the replies identified by the clinicians as either facilitative or deleterious. The total score on the SIRI-R represents the number of correct responses the participant identified. To be able to add up the correct replies, participants’ responses were recoded prior to analysis. Participants’ replies that correctly identified the Helper response were recoded and scored as 1 whereas incorrect replies were recoded as 0. For example, Helper A’s response was considered facilitative to intervention and thus, it was assigned by the participant a value of +1 to +3, was recoded as correct and recoded as 1. Helper B’s response was considered deleterious to intervention and if correctly identified by the participant with a value of -1 to -3, was recoded as 1. Helpers’ responses assigned the value of zero were scored as incorrect and recoded as 0. A participant’s total score consisted of the number of statements correctly identified. The maximum possible score for the six statements was twelve. Hence, a higher score reflects more awareness of appropriate suicide intervention skills.

All of the statistical analyses were performed with SPSS® version 16.0 for Windows®.

**Ethical Considerations for the Program Evaluation**

Ethics approval to conduct the program evaluation was sought and obtained from the Research Ethics Board at St. Michael’s Hospital (SMH). Workshop participants were made aware that their participation in the program evaluation was entirely voluntary, and that they may at any time, by not completing subsequent questionnaires, withdraw from the program evaluation.
without any effect on future care they received at SMH or their employment at the TTC. All of the participants were also informed that they could feel free not to answer any question which made them feel uncomfortable, and that there was no right or wrong answers to the questions.

All participants in the qualitative evaluation signed a detailed informed consent form prior to commencing the interviews (Appendix L). A copy of their signed consent form was given to all participants. Participants were informed in writing and orally that their participation in the program evaluation is voluntary and that they may at any time, withdraw from the program evaluation without any affect on the care they receive at SMH or their employment at the TTC.

Privacy and confidentiality of participants and data was ensured through a variety of means: all data, tapes, and transcripts had all identifiers removed and a program evaluation identification number was assigned; the audio-taped recordings were destroyed upon completion of the study; the program evaluation data was kept secured and separate from clinical files; access to the data was restricted to myself and researchers involved in the program evaluation. When referencing participants in the findings, their occupation and their assigned identification number are used instead of their names and gender to provide further anonymity. The measures taken to ensure confidentiality were explained to participants in the both the quantitative and qualitative studies. Participants in the qualitative study were informed about privacy and confidentiality measures taken prior to obtaining consent, while participants in the quantitative study were informed at the start of the workshop at the time the questionnaires were distributed. Participants were also assured that their non participation would not affect any treatments they receive now or in the future.

It was not anticipated that the interview evoked any emotional distress or acute discomfort in the qualitative evaluation participants. None-the-less, at the end of each interview, participants were given a list of community resources they may contact in the event they felt triggered or distressed as a result of the interview (Appendix M).

**Phase Two: TKM Level I- Reaction & TKM Level III - Behaviour**

**Qualitative Evaluation**

The overall goal of phase two of the program evaluation was circumscribed in scope and purpose. Phase two aimed at complementing the quantitative study carried out in phase one by adding depth and breadth to the gatekeeper program evaluation.
Phase two used a mixed-method design that combined a dominant qualitative method with a less-dominant quantitative method (Creswell, 2003; Teddlie & Tashakkori, 2003). The smaller quantitative component (five questions pertaining to competencies) was integrated in the qualitative interview to evaluate statistically perceived competencies and link them to changes in behaviour in the workplace. The face-to-face, personal semi-structured interview was chosen as the interview mode in this research evaluation. The Interviews were attributed dominant method status. This does not imply that the resultant quantitative data were any less valuable. The quantitative and qualitative methods simply produced information that served different purposes — breadth versus depth — both of which were important in the research. The quantitative and qualitative data were collected concomitantly.

As previously stated, The Kirkpatrick Model (TKM) provided the analytical framework for this program evaluation; hence, it guided the data collection and analysis of phase two of the evaluation. Further details on how TKM was utilized in phase two of the program evaluation are described later in this chapter.

Face-to-face, semi-structured interviews were conducted with a purposive sample of TTC employees from among those who attended the safeTALK workshop and who had volunteered to participate in phase two of the evaluation. The sample was drawn from two key groups of TTC employees who were likely to be involved with distressed at-risk of suicide patrons i.e., TTC Special Constables and Transportation Supervisors. Content analysis (Downe-Wamboldt, 1992; Boyatzis, 1998) was chosen as the method of analysis as it was judged to be the most appropriate method to analyze the interviews to meet the objectives set forth for the program evaluation and concomitantly, to conform with the analytical framework which evaluates reactions, learning, behaviour, and results (Kirkpatrick, 1959a; 1959b). A further discussion on content analysis is provided later in this chapter.

To meet the purpose of further describing content analysis as well as to detail other components of phase two methods, this section of the chapter is organized as follows. First, the interview guides used to collect the data and how they were developed are described. Next, the study’s recruitment procedures are described. The data collection and analysis which includes interviewing and coding are subsequently detailed. Finally, measures used to promote and ensure the trustworthiness of the findings are discussed.
**The Interview Guides**

Face-to-face, personal, semi-structured interviews were conducted for the purpose of garnering the participants’ reactions to the suicide prevention program and their perceptions on the impact it had on their competencies to intervene with a distressed, at-risk of suicide patron. Additionally, it sought to capture participants’ experiences with transferring their newly acquired knowledge and skills to behaviours in the workplace.

Interviews are a source of meaning and elaboration for program observations (Patton, 2002). Interviews help elucidate subjective meaning, experiences, beliefs and attitudes and permit the researcher to enter into another person’s world, to understand another person’s viewpoint and add the inner perspective to outward behaviour (Britten, 1995; Johnson, 2002; Patton, 2002; Rubin & Rubin, 1995). Interviews allow “an authentic gaze into the soul of another” (Silverman, 1994).

Two separate, albeit similar, interview guides were developed for the qualitative component of phase two of the program evaluation. The interview guides for TTC Special Constables and Supervisors were developed in consultation with several professionals (i.e., Dr. Ramsey, Dr. Mishara, John O’Grady and members of the Program Advisory Committee). Several drafts incorporating the different comments offered by those who were consulted were produced. The final versions of the interview guides are enclosed in Appendix J.

The interview guides were used as *aides-memoir* to serve as memory prompts to remind the researcher to pursue topics to be addressed in the interview i.e., reactions to the training, knowledge and skills acquisition, and to facilitate the exploration of these topics with the interviewees; however, the researcher did not restrict herself to the questions sets outlined in the interview guides. The interview guides were used to create an open environment in which the interviewee could reflect on the topics introduced by the researcher within the context of their own experience and reactions to the *safeTALK* gatekeeper program (Ezzy, 2002, p. 11). While the interview guides helped keep the interaction focused, it allowed individual perspectives and experiences to emerge (Patton, 2002).

The semi-structured interview style allowed the interview to retain a high degree of flexibility according to the participant’s experience and encouraged open dialogue that extended beyond the parameters set by the interview guide (Bryman, 2001, p. 107). The interview questions consisted of open-ended questions that delineate the area to be explored, yet allowed
the interviewer and interviewee to diverge and focus on a topic in greater detail (Britten, 1995, p. 251)

The interview guides consisted of several short rating scales and open-ended questions that sought to evaluate participants’:

- perceived comfort, confidence and competence in intervening with suicidal patrons;
- perceived appropriateness of their interventions;
- level of satisfaction with the gatekeeper program

The interview guide for TTC Special Constables consisted of the following sections:

1. A short demographics face sheet
2. Experience with Mental Health Act apprehensions (MHA)
3. Training
4. Apprehensions and interventions
5. Stress and coping.

The interview guide for subway supervisors varied slightly from the aforementioned and consisted of the following sections:

1. A short demographics face sheet
2. Experience with suicidal patrons
3. Training
4. Interventions
5. Stress and coping

The data collected from the stress and coping questions are related to an ongoing study; thus, were not included in the program evaluation.

**Participants Recruitment**

Sampling for the qualitative component was purposeful. Volunteers were sought from among the TTC Special Constables and Supervisors who had attended the *safeTALK* workshops. All Special Constables and TTC Supervisors who had attended the suicide prevention training were informed by their respective Superintendents of the qualitative study component of the program evaluation and were asked to volunteer and partake in the qualitative interview. Employees who were willing to participate contacted this researcher directly and scheduled an appointment for the interview. These two occupational groups were selected from among the
workshops attendees as they represented the two occupational groups called upon to deal with distressed patrons, and would have had the opportunity to apply the newly acquired knowledge and skills to situations in the workplace. It was thus deemed they would be in a position to provide the richest information and insight into participants’ reactions to the gatekeeper program and experiences with transferring newly acquired skills to situations in the workplace (Patton, 2002, pp. 230-242).

Sample Characteristics

The sample for the study consisted of TTC employees who had attended the workshops offered in the latter part of 2005. A total of 13 workshops were offered between September 30th and December 31st, 2005. A total of 176 employees attended the workshops; of those, 32 were TTC Special Constables and 75 were Transportation Supervisors (subway and surface). As indicated in Table 4, of the 176 employees who attended the safeTALK workshop, 136 employees participated in the program evaluation; of these, 22 were TTC Special Constables and 58 were Transportation Supervisors. In total, 30 employees were interviewed: 17 TTC Special Constables and 13 Transportation Supervisors. As indicted in Table 8, 16 of the TTC Special Constables and 12 of the Supervisors were males (93.3%). The average age of participants was 45.8 with 15.4 years of tenure at the TTC.

Data Collection and Analysis

Design and Procedures

The aim and purpose of data collection in qualitative research is to gather descriptions of the research participants’ lived experiences (Britten, 1995; Charmaz, 1991, p. 385; Johnson, 2002; Patton, 2002; Rubin & Rubin, 1995). The data in qualitative research can be broadly described as human experiences that are expressed and recorded in various forms. The process of data collection is an act performed by the researcher to gain and develop a deeper understanding and explore the lived experiences of the research participants. (Britten, 1995; Charmaz, 1991, p. 385; Patton, 2002; Johnson, 2002; Rubin & Rubin, 1995).
**Interviewing**

As noted earlier, the purpose of the qualitative interviews was to obtain participants’ perspectives and reflections on the *safeTALK* training effects on their knowledge, attitudes, and intervention abilities. Specifically, the qualitative study sought to evaluate participants’ perceived comfort, confidence and competence in intervening with suicidal patrons. The qualitative study also sought to obtain participants’ overall satisfaction with the suicide prevention program and participants’ thoughts and impressions of the training.

Researchers have viewed the qualitative interview as the democratization of the interpretations of interviewee’s experiences (Cisneros-Puebla, Faux, & Mey, 2004). The interview is a collaborative dialogue with the research participant in an attempt to accurately understand his/her experience. Since language is a reflective process, the interview is not merely a method to gather data about lived experience; it provides the research participant with an occasion to reflect on the experience. As noted earlier, the interviews conducted in this program evaluation were face-to-face and semi-structured. The interviews were conducted between July 2006 and the end of June 2007. All interviews were conducted by the researcher with each of the participants of phase two of the evaluation. Interviews were conducted sequentially. As potential participants contacted the researcher and verbally agreed to participate in the qualitative evaluation, an interview time and place were set.

Semi-structured interviews were conducted on the basis of a loose structure which consists of open-ended questions that delineate the area to be explored, yet allowed the interviewer and interviewee to diverge and focus on an idea in greater detail (Britten, 1995, p. 251). This interview process sought not to find facts through rigid questions-and-answer interviewing, but rather, to allow the interviewees to bring forth information as they saw fit. For this reason semi-structured, open-ended questions were phrased to the participants in order to allow them the freedom to voice their experiences and reactions to the suicide prevention program without being restricted by predetermined answers. The interview permitted for “a directed conversation that elicited inner views of respondents’ lives, experiences, and observations” (Charmaz, 1991, p. 385).

The interviews consisted of a short rating scale and open ended questions in order to evaluate the:

a. perceived appropriateness of their intervention;
b. perceived appropriateness of their referral plan;
c. perceived comfort, confidence and competence in intervening with suicidal patrons;
d. satisfaction with the gatekeeper program.

Written consents were obtained from participants prior to the start of the interview. Participants were informed at the start of the interview that there were no right or wrong answers and what the researcher sought was to elicit their impressions, reactions and experiences to the gatekeeper program. The researcher did not restrict herself to the question sets outlined in the interview guides. The interview questions were used as guides and there was a conscious effort to maintain flexibility and allow the participants to lead the interviews. Participants were informed at the start of the interviews that questions seeking additional clarification, details and understanding might be asked during the interview. At the completion of the question set, additional questions were asked. Some of these questions requested an expansion or clarification of what had been said in the interview. Additional questions were also based on statements made in earlier interviews. Comments were solicited on these topics if they did not come up in answers to the question sets.

Throughout the interview, the researcher was sensitive to the language and concepts used by the participants and kept the interview interactive and flexible (Britten, 1995, p. 252). Although the interviews were the first meetings between the participants and the researcher, the interviews were conversational with the initial stages of the interviews devoted to establishing rapport with the participants by creating a milieu of trust, empathy, understanding, and neutrality (Patton, 2002, p. 366). Interviews involve an interpersonal engagement within a relationship of safety and trust. Rapport is a salient component of optimal research interview (Patton, 2002, p. 366). Throughout the interviews the researcher adopted the stance of learner while the participants were regarded as the experts.

Most of the interviews lasted between 45-60 minutes and with the exception of one interview which was conducted at St. Michael’s Hospital, all other interviews were conducted during work hours in locations chosen by the participants: e.g. special constables’ division office, subway station control towers, the transit control office. The interview locations provided on the job contexts for the interviews and the individual location was recorded in the field notes.
written following each interview. The field notes consisted of short summaries that captured major elements and impressions of the interviews.

Although it was not anticipated that the interviews engendered emotional distress or acute discomfort in participants, at the end of each interview, study participants were given a list of community resources they could contact if they felt triggered or distressed as a result of the interview (Appendix I & J).

Table 8: Phase two Sample characteristics

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>TTC Special Constables</th>
<th>Transportation Supervisors</th>
<th>ALL phase two participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of participants</td>
<td>17</td>
<td>13</td>
<td>30</td>
</tr>
<tr>
<td>Age</td>
<td>43.6 (9.5)</td>
<td>48.6 (7.2)</td>
<td>45.8(8.8)</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>16 94.10%</td>
<td>12 92.30%</td>
<td>28 93.30%</td>
</tr>
<tr>
<td>Female</td>
<td>1 5.90%</td>
<td>1 7.70%</td>
<td>2 6.70%</td>
</tr>
<tr>
<td>Marital Status</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Married or cohabiting</td>
<td>15 88.20%</td>
<td>10 69.20%</td>
<td>25 83.30%</td>
</tr>
<tr>
<td>Separated</td>
<td>0 0%</td>
<td>2 15.40%</td>
<td>2 6.70%</td>
</tr>
<tr>
<td>Single</td>
<td>2 11.80%</td>
<td>1 7.70%</td>
<td>3 10.00%</td>
</tr>
<tr>
<td>Education</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High school</td>
<td>2 11.80%</td>
<td>4 30.80%</td>
<td>6 20.00%</td>
</tr>
<tr>
<td>Some college or university</td>
<td>2 11.80%</td>
<td>8 61.50%</td>
<td>10 33.30%</td>
</tr>
<tr>
<td>Completed college or university</td>
<td>13 76.50%</td>
<td>1 7.70%</td>
<td>14 46.70%</td>
</tr>
<tr>
<td>Years of employment at TTC</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>13.2 (8.2)</td>
<td>18.3 (8.0)</td>
<td>15.4 (8.4)</td>
</tr>
<tr>
<td>Years at the current job</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>10.8 (6.5)</td>
<td>8 (7)</td>
<td>9.6 (6.7)</td>
</tr>
</tbody>
</table>

**Data Entry**

All of the interviews were audio-taped in full (with consent). Each interview was recorded on a separate audiotape cassette. The audio-tape recordings were later transcribed verbatim by a
professional external transcriber. All interviews were transcribed verbatim to best represent the dynamic nature of the living conversation (Riessman, 1993; Seidman, 1991).

To verify that data have been transcribed to an appropriate level of detail, all of the transcriptions were later checked by the researcher for accuracy. The data management and analysis was aided by the use of the qualitative software QSR N6. The primary benefit to having used a qualitative software program was that it simplified the mechanical aspects of the data analysis and allowed the researcher to be more effective and thorough; however,

“The thinking, judging, deciding, interpreting, etc. are still done by the researcher. The computer does not make conceptual decisions, such as which words or theme are important to focus on, or which analytical step to take next” (Teesh, 1991, pp. 25-26).

**Data Analysis**

Data analysis in quantitative and qualitative research typically includes one or more approaches as depending on the nature of the data and the research questions. In qualitative research there are a variety of established procedures for analyzing data (Huberman & Miles, 1994). Generally, these procedures involve converting raw narrative data (e.g., field notes, audiotapes) into partially processed data (e.g., transcripts, which are then coded and subjected to any one of a number of analysis schemes such as theme analysis or constant comparative analysis).

Qualitative research, based on data from interviews requires understanding and cooperation between researcher and participants such that text based on interviews is mutual, contextual and value bound (Lincoln & Guba, 1985). The assumption is that text always involves multiple meanings and there is always a degree of interpretation when approaching text. Qualitative analysis requires some creativity, for the challenge is to place the raw data into logical, meaningful categories; to examine them in a holistic fashion; and to find a way to communicate this interpretation to others (Patton, 1990a; 1990b). Bogdan and Biklen (1982) define qualitative data analysis as:

working with data, organizing it, breaking it into manageable units, synthesizing it, searching for patterns, discovering what is important and what is to be learned, and deciding what you will tell others (p.145).
The analysis in this evaluation drew on two primary sources: 1. evaluation questions that were generated during the conceptual and design phase of the evaluation; 2. analytical insights and interpretation that emerged from data collection.

Content analysis has been described as: the systematic examination of text by identifying and grouping themes and coding, classifying, and developing categories (Pope & Mayes, 1995). Content analysis is a process for systematically analyzing messages in any type of communication and it concerns itself with meanings, intentions, consequences, and contexts. “Text based on interviews is shaped within and by the interaction between the researcher and the participants and can be seen as a communication act” (Graneheim & Lundman, 2004, p. 111) and each “communication act has a content aspect and a relational aspect and the latter classifies the former” (Watzlawick, Bavelas, & Jackson, 1967, p. 54). The goal of content analysis is to enhance the inferential quality of the results by relating the categories to the context that produced the data. Content analysis involves identification and analysis of themes and patterns of similarity within qualitative research (Braun & Clark, 2006). It is not grounded in any particular theoretical and epistemological framework and can therefore be applied across a wide range of qualitative research approaches (Braun & Clark, 2006). In seeking to analyze data, content analysis can either identify themes pertaining to a particular research question (deductive analysis) or it can identify themes that are observed across the entire data range (inductive analysis) (Braun & Clark, 2006). One characteristic of qualitative content analysis is that method, to a great extent focuses on the subject and context and emphasizes differences between similarities within codes and categories. This process can be conceptualized as a process of “mapping” (Miles & Huberman, 1994).

A fundamental issue when performing qualitative content analysis is to decide whether the analysis should focus on manifest or latent content. The analysis of what the text says deals with the content aspect and describes the visible, obvious components referred to as manifest content (Downe-Wamboldt, 1992; Kondracki, Wellman, & Amundson, 2002). In contrast, analysis of what the text talks about deals with the relationship aspect and involves an interpretation of the underlying meaning of the text, referred to as the latent content (Downe-Wamboldt, 1992; Kondracki et al., 2002). Both manifest and latent content deal with interpretation, but the interpretation varies in depth and level of abstraction.
The method of analysis chosen for this study was a hybrid approach of qualitative methods of content analysis and it incorporated both the data-driven inductive approach and the deductive a priori template of code approach outlined by Crabtree and Miller (1999). The hybrid approach was used to complement the research questions by combining the process of deductive (manifest) content analysis with themes that emerge directly from the data using inductive (latent) coding.

Patton (1987) asserted that there is no precise point at which data collection ends and analysis begins, nor are analysis and interpretation precisely separated. In the course of data collection, ideas about analysis and interpretation emerge. The analytical approach commences during data collection and continues after the data collection ceases. Overlapping of data collection and analysis improves both the quality of data collected and the quality of analysis with the caveat that the researcher is attentive and does not allow early interpretation to bias additional data collection.

Constructing a particular analytical process depends primarily on the goals of the research and analysis. This analytical process was guided by The Kirkpatrick Model (TKM). The analytical framework, which guided the two phases of the program evaluation, provided the “structure, the scaffolding, the frame to a study or program evaluation” (Merriam, 1998, p. 45). The a priori categories involved developing a coding template which summarized broader categories identified by the theoretical framework and research questions; thus, the analytical process was more structured and used a priori codes in a template that encompassed content areas that corresponded to the levels of analysis identified in the theoretical framework i.e. reaction, learning, and behaviour.

Template analysis refers to a method of thematically analyzing qualitative data. Template analysis involves developing a coding template which summarizes themes identified by the researcher as important in a data set, and organizes them in meaningful and useful manner. The use of a “template analysis” (Crabtree & Miller, 1999) allowed for a more focused and time efficient analysis and permitted the researcher to focus on particular aspects of the text without delving into an intense line-by-line analysis until after coded segments had been identified and sorted. The template process reduced the amount of data being considered at any one time and brought together related segments of the text earlier in the process, which facilitated making the links and connections.
The choice of a code manual for the study was important, because it served as a data management tool for organizing segments of similar or related text to assist in interpretation (Crabtree & Miller, 1999; King, 2004). The template was developed a priori based on TKM which served as the analytical framework for the program evaluation. Three broad categories i.e., Reactions, Learning, and Behaviours formed the code book. An essential step in the development of a useful template for analysis is to determine the applicability of the code to the raw data (Boyatzis, 1998). The initial template was applied to the whole data set and it was modified in light of careful consideration of each transcript. Once the final version was defined and all the transcripts had been coded to it, the template served as a basis for interpretation or illuminating the data set.

One of the most basic decisions when using content analysis is selecting the unit of analysis (Downe-Wamboldt, 1992). In this analysis it was considered most suitable to consider the whole interview as the unit of analysis, while the meaning unit was the constellation of words or sentences or paragraphs that contained aspects related to each other through their content and context (Downe-Wamboldt, 1992).

Coding

Using template analysis (Crabtree & Miller, 1999; King, 2004), the transcripts were coded into broad themes based on the research objectives and interview questions created to evaluate the gatekeeper program which corresponded to the TKM levels: reaction, learning, and behaviour.

The interviews were read through several times to obtain a sense of the whole. Next, the text about the participants’ experiences of attending the workshops was extracted and brought together into one text, which constituted the unit of analysis. The text was then divided into meaning units that were later condensed. The condensed units were abstracted and labeled with a code. The whole context was considered when condensing and labeling meaning with unit codes. The various codes were compared based on differences and similarities and sorted into subcategories and categories, which constituted the manifest content. The coding encompassed successively narrower, more specific ones. Where such segments corresponded to a priori themes, they were coded as such; otherwise new themes were defined to include the relevant
material and were organized into the initial template after reading through three of the fifteen transcripts in the study.

The labeling of meaning units has been referred to as a code. Open coding is defined as “the analytical process through which concepts are identified and their properties and dimension are discovered in the data” (Strauss & Corbin, 1998, p. 101); thus the primary goals of open coding are to conceptualize and categorize data. Although a priori categories were dictated by the theoretical framework; nonetheless, individual concepts were gathered to form abstract categories and subcategories using words elicited by the participants themselves, what Strauss and Corbin (1998) referred to as “in vivo” language. Using “in-vivo” codes reduced the potential for researcher bias in labeling and developing concepts.

The template initially provided the analytical tool to handle copious amounts of raw data, and allow the researcher to be systematic while the open coding assisted in considering alternative meaning and concomitantly to identify and develop related concepts. The template was further modified and developed and each theme was subjected to the formation of a more thorough analysis and to more specific categories within each theme. The hierarchical coding allowed the researcher to analyze texts at different levels of specificity. Broad higher-order codes helped provide a general overview of the direction of the interviews, while detailed lower-order codes enabled fine distinctions to be made, both within and between participants (Crabtree & Miller 1999; King, 2004).

Creating categories is the core feature of qualitative content analysis. A category is a group of content elements that shares commonality. Patton (1987) emphasizes that categories are internally homogeneous and externally heterogeneous. Categories must also be exhaustive and mutually exclusive. A category refers to a descriptive level of content of the text and is an expression of a manifest content of the text. Categories often include a number of sub-categories at varying levels of abstraction. Themes were created by linking meanings together; reducing meaning units, codes and categories, on an interpretive level. A theme is the expression of the latent content of the text (Downe-Wamboldt, 1992).

The memos served as a document of the researcher’s mind processes, bases for decisions, and constructions of meaning within the research process and research participants’ relationship. The memos were operational, observational personal and/or analytical. Post-interview memos recorded the non-verbal aspects of the interview that were not evident from the audiotape of the
interview, such as body language, affect, temperament, attitudes, and environmental distractions. Memos written throughout the evaluation process also provided the opportunity for reflexivity and recorded the researcher’s thoughts, ideas, insights, and reflections. When necessary, data memos were created and attached to specific lines of data within interviews to explain the meaning of the text. Analytical memos were written throughout the analytical process to record and trace conceptual thoughts and ideas and illuminate interrelated processes and help link analytical interpretation with empirical reality.

**Trustworthiness**

Research findings should be as trustworthy as possible and every research study must be evaluated in relation to the procedures used to generate the findings. Within the tradition of content analysis, use of concepts related to the qualitative tradition, such as validity, reliability, and generalizability is still common (Downe-Wamboldt, 1992; Sheilds & King, 2001). The purpose of validity and reliability irrespective of the research tradition, quantitative or qualitative, have essentially the same meaning (Long & Johnson, 2000, p. 31) and “all aspects of trustworthiness are intertwined and interrelated’ (Graneheim & Lundman, 2004; 109)

Qualitative research is circumscribed by the context in which the research is conducted. Any results and conclusions are “true” only to that particular setting and thus not generalized. Results from qualitative studies; however, increase and augment our knowledge of subjective experiences. Whilst bias of any kind is rigorously guarded against in any scientific research, it is inevitable in qualitative research. In any qualitative study, the researcher is the primary instrument for the research process (Guba & Lincoln, 1981, Patton, 1987); thus it is inevitable that the researcher will bring to the research process her attitudes, biases, skills, and knowledge which subsequently influence the collection, selection, and interpretation of data (Broom, 2005, p. 72). Furthermore, the researcher’s behaviour affects participants’ responses, and thus affects the directions of findings (Finlay, 2002). The qualitative research process itself has the potential to alter the phenomenon being studied. Research is regarded as “a joint product of the participants, the researcher, and their relationship...” (Finlay, 2002, p. 531). Interview data is ultimately co-constructed by the interviewer and the interviewee, a dialogue between two participants, rather than an objective account of a particular event. Consequently, it is incumbent on the researcher to find ways to analyze how their own subjectivity, personal biography, and
biases influence their research. Although researcher’s bias is assumed within the qualitative paradigm, the researcher was vigilant to bracket personal experiences and suppositions and biases she may have had to allow the research process to develop and unfold as naturally as possible. The following discussion presents this researcher’s biases prior to conducting the program evaluation. Before undertaking the program evaluation as a doctoral project, this researcher was involved in the suicide prevention initiative phases and the implementation of the gatekeeper program at the TTC. For the purpose of developing the list of warning signs that were later incorporated in the training this researcher interviewed numerous TTC employees and was involved in the development of written materials on suicide distributed to TTC employees. This researcher was also consulted by LivingWorks education in the development of the workshop content and the video vignettes incorporated in the training. There were no assumptions made by the researcher regarding the outcomes of the program evaluations, nor did the researcher have any stakes in the outcome of the evaluation beyond the program evaluation constituting the researcher’s doctoral dissertation.

The concepts of credibility, dependability, and transferability have been used to describe various aspects of trustworthiness (Lincoln & Guba, 1985; Patton, 1987). The following strategies were employed to promote and demonstrate trustworthiness and rigour: methodological coherence i.e., congruence between research questions, data collection, analysis and interpretation; sampling adequacy i.e., recruitment strategies which led to data saturation i.e. collecting additional data seemed counterproductive and the data became repetitive and redundant (Lincoln & Guba, 1985, p. 202; Morse & Field, 1995, p. 147; Strauss & Corbin, 1998, p. 136). Additionally, data collection and analysis were systematic and transparent. Establishing the trustworthiness of the transcripts is a fundamental component of rigour in qualitative research. Special efforts were made to ensure the quality and accuracy of the transcripts. Each transcript was verified prior to undertaking the analysis of the textual data. The researcher returned to the data several times to verify that concepts, categories, explanations and interpretations made sense and reflected the data.

Credibility deals with the focus of the research and refers to confidence in how well data and process analysis address the intended focus. It deals with the focus of the study, the selection of context, participants, and approach to gathering data. Choosing of participants with various experiences increases the possibility of illuminating the research questions from a variety of
perspectives (Patton, 1987). In this evaluation, participants of various ages, and occupational
groups and work experience provided various perspectives and contributed to a richer variation
of the evaluation. Selecting the most appropriate method of data collection was an essential
factor in establishing credibility. The participants’ reflections, conveyed in their own words,
strengthened the face validity and credibility of the research (Patton, 2002). Choosing a mixed-
methods approach and using more than one data collection approach permitted the researcher to
combine strengths and correct some of the deficiencies of any one source of data and allowed the
researcher to capture the intricacies of a program evaluation. Additionally, the approach allowed
for triangulation of qualitative and quantitative data; thus, potentially enhancing the
trustworthiness of the evaluation findings.
CHAPTER FIVE
QUANTITATIVE RESULTS

The results of the quantitative component of the study are presented in this chapter. The results focus on the research questions related to the immediate and long-term training effects on study participants’ attitudes, knowledge, and intervention skills, and examines the training effects on the various occupational groups.

The results of the training effects are discussed under headings related to the study hypotheses and are linked to the TKM, the analytical framework used to guide the evaluation: i.e.: Reactions, Learning, Behaviours, and Results. The extent to which study participants acquired knowledge, changed attitudes, and/or increased or enhanced intervention skills as a result of attending the safeTALK and suicideAWARE suicide prevention programs corresponds to Level 2 (Learning) of the analytical framework.

The chapter commences with a brief description of the sample characteristics and the study’s participants’ experience with suicide and Mental Health Act apprehensions. The results of the applicable analyses are presented sequentially with the analyses organized in separate sub-sections following the order of the hypotheses: Hypothesis 1: Knowledge, Hypothesis 2: Attitudes, and Hypothesis 3: Intervention knowledge and skills. In each of the sub-sections, descriptive statistics will be presented initially for each of the outcome measures examined in the present study. These results are followed by comparisons of the linked and unlinked data statistics.

Missing data are common problem in empirical research and occur essentially in every study. Failure to capture all data for all of study measures presents serious threat to internal validity. There are various methods of handling missing data. Simple and frequently used methods include the missing-indicator analysis, overall mean imputation, and complete case analysis (Donders, van der Heijden, Stijne, & Moons, 2006). Discarding incomplete cases, known as complete case analysis, has several disadvantages that affect the strength, integrity, reliability, and/or validity of causal inference, and affect internal validity of the findings. First, the final sample may form a subgroup that inadequately represents the study participants. Second, the complete cases (i.e., linked data) and incomplete cases (i.e., unlinked data) groups may have different attributes that serve as confounds, and thereby threaten the internal validity.
of the findings. Thirdly, unbalanced cell size may limit the use of some statistical tests and may reduce statistical power to detect significant training effects. Differential attrition may also result in unequal pre-training vs. post-training group size. Unequal cell size may result in violation of assumptions, normality and homogeneity of variance, that are the basis of F-test. Missing data also influence the magnitude of effect size through either artificial inflation from uncontrolled extraneous factors or attenuation from increase with-in group variance.

To examine whether the missing data in this study inadvertently introduced any systematic bias or threat to the internal validity of the study, statistical analyses were conducted with linked (i.e., complete cases) and unlinked (incomplete cases) data collected throughout the study. While utilizing unlinked data has increased the statistical power to detect significant training effects, the comparison to the linked data analyses addressed concerns about possible biases in the unlinked data set. The results of these comparative analyses are presented alongside findings from the Linear Mixed-effects Models (LMEs) analyses.

The results of the Linear Mixed Model analyses are followed by post hoc contrast analyses. Finally, in each of the sub-sections the results of separate analyses performed for the three occupational groups of interest, i.e., Special Constables, Transportation Supervisors, and Train Operators are presented. The presentation of the occupational groups’ results follows the aforementioned sub-section organization and presentation of results.

**Significant Sample Characteristics**

The *safeTALK* training workshops were offered to all Transportation Supervisors, Special Constables, Trainers, Safety and Hygiene personnel. Attendance in the *safeTALK* workshops was voluntary. The *suicideAWARE* training workshops were offered to all train operators as part of their bi-annual recertification program. During the study period 360 employees attended the *safeTALK* and *suicideAWARE* suicide prevention workshops. Of the 360 TTC employees who have attended the suicide prevention workshop 85.8% (309/360) participated in the program evaluation study. Follow-up questionnaires were mailed to 307 participants and 174 questionnaires were completed and returned. The effective follow-up response rate; therefore, was 56.7% (174/307).
Of the 309 study participants 63 safeTALK and 65 suicideAWARE, participants indicated their TTC employee number on the completed pre-, post-, and follow-up questionnaires. This allowed for their responses to be linked for complete case analysis.

Key demographic and occupational characteristics of the study participants are presented in Table 9. The majority of study participants was male (86.6%), over the age of 35 (83.7%), and nearly half of the sample (41.7%) was employed by the TTC for over 10 years, with the majority of study participants (55.5%) employed as Train Operators. There were no statistical differences in demographics between safeTALK and suicideAWARE participants. The sole difference between the two groups of workshop participants was their occupation. All of the suicideAWARE participants were Train Operators while the safeTALK attendees consisted of Special Constables, Transportation Supervisors, and various other TTC personnel (trainers, occupational health, and safety personnel).

All of the Special Constables who participated in the study were males (100%), over the age of 35 (77.3%), and were employed by the TTC for less than 10 years (81.8%). Among the Transportation Supervisors the majority (86.2%) were male, over the age of 35 (88.2%) with less than 10 years experience as Transportation Supervisors (84.5%). Likewise, the group that consisted of trainers, and various other personnel were predominantly males (64.9%), over the age of 35 (93%), and with the majority (74%) reporting less than 10 years of experience working at the TTC. A similar trend was observed among Train Operators who had attended the suicideAWARE workshop and participated in the study: the vast majority was male (92.3%), over the age of 35 (81.7%) with less than 10 years experience as Train Operators (82.7%).

Experience with Suicide and Suicidal Behaviours

Contact with suicide and suicidal behaviour varied in scope and was dependent on the occupational role the participant had at the TTC. TTC Special Constables’ involvement with suicide and suicidal behaviour was diverse and ranged from taking a patron into custody if risk of self harm was suspected, to photographing suicide scenes. Likewise, Transportation Supervisors were dispatched by Transit Control to a subway station when the behaviour of a patron on a subway platform signaled distress and Transportation Supervisors were directly involved in intervening with a distressed patron.
Additionally, Transportation Supervisors were also dispatched to the scene of a suicide and were responsible for the removal of the body or body parts from under the train, cleaning up of the tracks, and ensuring service resumed swiftly with minimal delays. Conversely, the involvement of Train Operators was limited to either driving the train or being on guard duty when the train was involved in a “priority one” incident, i.e.: suicide or suicide attempt. The question in the pre-training questionnaire was broad and non-specific and did not define the term “contact”; thus the responses encompassed any involvement with suicide and suicidal behaviour from intervention to the cleanup of the suicide scene.
The extent of the participants’ experience with suicide and suicidal behaviour during their years of employment at the TTC is summarized in Table 10. The study participants’ experience with suicide and suicidal behaviour varied and the reported number of incidences during all their years of employment at the TTC ranged from 0 to 70 with an overall mean number of reported incidences of 2.8 ($SD = 7.7$, $Mdn. = 0$). Over half of the participants (51.6%) reported no experience with suicidal behaviour, two-fifths (40.5%) had been involved in less than 10 incidents; however, one study participant reported involvement in as many as 70 suicidal behaviour-related incidents. The vast majority of study participants (86.4%) had no involvement with a suicide-related incident in the year prior to attending the suicide prevention training.

**Experience with Mental Health Act Apprehensions**

Section 17 of the Ontario Mental Health Act (MHA) gives TTC Special Constables the power to take a patron into custody and transport them to a hospital for a psychiatric assessment if the Special Constable has reasonable and probable grounds to believe that the patron acted or is acting, on TTC property, in a disorderly manner (i.e.: behaviour that is to some extent irrational and not necessarily unruly). With regards to suicidality, the Special Constable must have reasonable cause to believe that the patron has threatened or is threatening to cause bodily harm to self or has attempted or is attempting to cause bodily harm to self. The Special Constable must be of the opinion that the person is apparently suffering from a mental disorder that will likely result in one or more of the following: serious bodily harm to self or serious bodily harm to others or serious physical impairment to self. (Bill 68, Ontario Mental Health Reform2000::http://www.ontla.on/web/bills/bills.detail.do?locale=en&BillID=576&Current=false 6/4/2010).

As Table 11 illustrates, the majority of safeTALK participants (63.8%) reported no experience with Mental Health Act (MHA) apprehensions; however, among those study participants who had reported experience with MHA apprehensions, their involvement ranged from 0 to 70 MHA apprehensions. Nearly one-quarter of study participants (23.9%) reported involvement in 1 to 10 apprehensions, and one in eight participants (12.3%) reported involvement in 12 to 70 apprehensions. The mean number of MHA apprehensions for all the years of employment at TTC was 2.43 ($SD = 7.6$, $Mdn. = 0$).
Table 10: Workshop Participants’ Experience with Suicidal Behaviour While with the TTC

<table>
<thead>
<tr>
<th>Time Frame of Experience</th>
<th>Experience</th>
<th>N</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>In all the years with TTC</td>
<td>No experience</td>
<td>158</td>
<td>51.60%</td>
</tr>
<tr>
<td></td>
<td>≤ 10 experiences (1-10)</td>
<td>124</td>
<td>40.50%</td>
</tr>
<tr>
<td></td>
<td>&gt; 10 experiences (12-70)</td>
<td>24</td>
<td>7.80%</td>
</tr>
<tr>
<td>In the past year</td>
<td>No experience</td>
<td>266</td>
<td>86.40%</td>
</tr>
<tr>
<td></td>
<td>≤ 10 experiences (1-7)</td>
<td>39</td>
<td>12.70%</td>
</tr>
<tr>
<td></td>
<td>&gt; 10 experiences (15)</td>
<td>3</td>
<td>0.97%</td>
</tr>
<tr>
<td>In the past month</td>
<td>No experience</td>
<td>294</td>
<td>96.70%</td>
</tr>
<tr>
<td></td>
<td>≤ 10 experiences (1-10)</td>
<td>10</td>
<td>3.30%</td>
</tr>
</tbody>
</table>

In the year prior to attending the workshop over three-quarters (78.7%) of the study participants had no involvement in MHA apprehensions and the mean number of reported MHA apprehensions was .79 (SD = 2.1 Mdn. = 0, range of 15 from 0 to 15). The number of study participants with no involvement in MHA apprehensions was greater in the month prior to attending the workshop, with the vast majority of the participants (93.4%) reporting no involvement in an MHA apprehension.

**Occupational Group and Experience with MHA Apprehensions and Suicidal Behaviour**

Among the Special Constables, their experience with suicide-risk MHA apprehensions varied and ranged from 0 to 70 incidents. In all their years of employment at the TTC, over two-fifths of the Special Constables (41.1%) reported less than 10 suicide-risk MHA apprehensions, whereas 58.8% reported having been involved in 12 to 70 suicide-risk MHA apprehensions. The
The overall mean number of suicide-risk MHA apprehensions was 25.8 ($SD = 25.7$, $Mdn. = 20$). Transportation Supervisors did not have the powers to apprehend patrons who were distressed and at risk of suicide and reported no involvement with suicide-risk MHA apprehensions.

Examining the data by occupations revealed that Special Constables experience with suicide ranged between 0 to 70 incidences. Nearly two-thirds of the Special Constables (58.3%) had been involved in more than 10 suicides throughout their period of employment at the TTC: in contrast, 11.8% had never been involved in a suicide incident. The overall mean number of suicide and suicidal behaviour incidences in which Special Constables reported involvement was 26.7 ($SD = 19.2$, $Mdn. = 21.0$).

Compared to Special Constables, Transportation Supervisors reported fewer experiences with suicide and suicide attempts. During all their years of employment at the TTC, Transportation Supervisors’ experiences with suicide and suicide attempts ranged from 0 to 50 incidences. Nearly one-quarter (23.1%) of the Transportation Supervisors had no experience with a suicide and almost a third (30.8%) had no experience with a suicide attempt. Over a third of Transportation Supervisors (38.5%) had been involved in less than 10 suicide incidences, and over half (53.9%) had been involved in incidences of attempted suicides. The overall mean number of involvement in suicide incidents was 15.2 ($SD = 19.5$, $Mdn. = 5$), while the overall mean number of involvement in attempted suicide incidents was 7.2 ($SD = 13.6$, $Mdn. = 2$).

Train Operators’ experiences with suicides and suicide attempts implied they had driven or had been on guard duty on a train involved in a suicide or a suicide attempt of a TTC patron. Of the 171 Train Operators who had participated in the study, nearly half (48.5%) reported no involvement in a suicide or a suicide attempt in all years of employment at the TTC. Among those Train Operators who had reported having had the experience of a patron’s suicide or a suicide attempt, 49.1% had been involved in 1 to 5 suicides, while 2.4% experienced 13 to 16 suicides. For Train Operators the overall mean number of suicides was 1.0 ($SD = 2.0$, $Mdn. = 1$). There were no statistically significant differences between the linked and unlinked data with regards to experiences with MHA apprehensions or suicidal behaviours.

The first study hypothesis predicted factual knowledge about suicide and suicide risk factors would increase after the training and the gains made in knowledge would be maintained over time. The Suicide Procedural Questionnaire (SPQ) and the Intervention Knowledge Test (IKT) were used to test the knowledge acquisition.
Table 11: Workshop’s Participants’ Experience with MHA Apprehensions While with the TTC

<table>
<thead>
<tr>
<th>Experience</th>
<th>N</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>In all the years with TTC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No experience</td>
<td>88</td>
<td>63.80%</td>
</tr>
<tr>
<td>≤ 10 experiences (1-10)</td>
<td>33</td>
<td>23.90%</td>
</tr>
<tr>
<td>&gt; 10 experiences (12-70)</td>
<td>17</td>
<td>12.30%</td>
</tr>
<tr>
<td>In the past year</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No experience</td>
<td>107</td>
<td>78.70%</td>
</tr>
<tr>
<td>≤ 10 experiences (1-10)</td>
<td>28</td>
<td>20.60%</td>
</tr>
<tr>
<td>&gt; 10 experiences (15)</td>
<td>1</td>
<td>0.70%</td>
</tr>
<tr>
<td>In the past month</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No experience</td>
<td>127</td>
<td>93.40%</td>
</tr>
<tr>
<td>≤ 10 experiences (1-2)</td>
<td>9</td>
<td>6.60%</td>
</tr>
</tbody>
</table>

**Suicide-risk Procedures Questionnaire (SPQ)**

One of the study’s objectives was to determine whether study participants’ knowledge of suicide-risk warning signs and the concomitant procedural actions that were need to be initiated changed as a result of attending the suicide prevention workshops i.e.: *safeTALK* or *suicideAWARE*. The knowledge was assessed using the Suicide-risk Procedures Questionnaire.

The Suicide-risk Procedures Questionnaire consisted of four multiple-choice questions describing behaviours that may signal distress and alert the observer to potential suicide risk. A participant total score consisted of the number of correct responses. The maximum possible score for the SPQ was four. For the purpose of the GEE analysis, two categories were created “≤3” and “=4” with the former responses coded as “0” and the latter coded as “1”. The categories were created using the pre-training median score of 4.
As illustrated in Table 12, prior to the training slightly over two-thirds of the participants (69.7%) had correctly answered all four questions in the Suicide-risk Procedures Questionnaire. Post-training, the proportion of participants who correctly answer all four questions increased from 69.7% to 83.2%. At the three-month follow-up the proportion of those who correctly answered all four questions increased to 88.1%. Based on the GEE, using an unstructured working correlation matrix, the increase in proportions of participants who responded correctly to all four questions post-training and at follow-up were statistically significant (p < .001).

Comparing the characteristics of participants who had low SPQ (≤ 3/4) scores (n = 206) with those who had high SPQ (4/4) scores (n = 88) suggested the two groups did not differ in demographic characteristics such as age (>35 85. vs. 83%, respectively), gender (87% males), years of employment at the TTC (< 10 years 82% vs. 83%), occupational group membership (Special Constables 8% vs. 9%; Transportation Supervisors: 16% vs. 20%; Train Operators: 53% vs. 52%), experience with suicide and suicidal behaviour (0 experience: 47% vs. 54%) and MHA apprehensions (0 experience: 66% vs. 65%).

An examination of the linked and unlinked scores indicated that there were no statistically significant differences between the proportion of participants who correctly answered all four SPQ questions in the linked and unlinked data. Likewise, no statistical significant differences were observed at the three-month follow-up between the data sets (see Table A in Appendix N).

As all of the scores were continuously distributed responses, linear mixed-effects models (LMEs) were used to describe the immediate and long-term effects of the training on knowledge acquisition and in following sub-sections on attitudinal changes and skill enhancement.

The LME approach allowed hypothesis testing and provided a flexible yet parsimonious way of modeling the association among repeated measurements and allowed for missing data. The simple fixed-effects structure included a three-level factor representing time (independent variable). Prior to fitting these models, preliminary analyses were conducted to verify the appropriateness of the approach. As the design comprised only three time points, there was limited choice of error structures for the residuals. A comparison of goodness of fit using Akaike’s Information Criteria and Schwartz’s Bayesian Criteria of two structures (compound symmetry and unstructured error structure variance-covariance matrix) was conducted to select the best correlation structure. To assess the differential, immediate and long-term effects of the
training on the SPQ mean scores, a linear fixed-effects model was performed using CS
correlation structure. As indicated in Table 13, the estimated marginal mean post-training
increased from 3.56 to 3.84 and was maintained at the three-month follow-up (M = 3.56 vs. M =
3.85). The analysis of variance showed that the effects of the training on the marginal estimated
means scores were significant (F\textsubscript{2,507} = 17.135, p < 0.001). Further post-hoc contrast analyses
using Sidak procedures indicated pre-post and pre-follow-up pairwise comparisons were
statistically significant (p < 0.001). Correspondingly, as illustrated in Table 14, there was a
statistically insignificant increase in the post-training SPQ mean raw score when compared to the
pre-training mean score (M = 3.61 vs. M = 3.78, p <.001) with an absolute change in the mean
score of +.17 and a relative increase in the post-training mean score of 7.2%. To contextualize
the effects of the training on knowledge, the magnitude of training effects was computed with
Cohen’s d for repeated measures. The formula used for the computation of the effect size
reported can be found in Appendix K. Cohen’s d test for repeated measures indicated a small
effect size corresponding to .41. Likewise, at the three-month follow-up there was a statistically
significant increase in the mean score (M = 3.61 vs. M = 3.87, p < .001) and a medium effect
size was maintained (Cohen’s d = .48). The absolute change in the mean score of +.26, and a
relative increase in the SPQ follow-up mean score of 4.7%.

**Linked vs. Unlinked Scores**

Comparisons of the linked and unlinked SPQ data using one-way ANOVA tests indicated
a statistical difference between the mean linked and unlinked scores (p = .05); however, further
post-hoc analysis using Sidak procedures comparing linked and unlinked pre-, post- and follow-
up mean scores indicated the differences were not statistically significant (p =.52, p = .31, p =
.63, respectively). Table A in Appendix N presents the results of mixed linear fixed-effects
analyses and the post hoc pairwise comparisons of the linked and unlinked data.

**SPQ and Occupational Groups**

As illustrated in Table 15, the proportions of Special Constables who post-training
correctly answered all four questions on the Suicide-risk procedural questionnaire increased
from 52.4% observed at pre-training to 100%. Although the numbers of Special constables who
correctly answered all four SPQ questions dropped slightly at the three-month follow-up to
90.9%, nonetheless, it remained high in comparison to the pre-training numbers.
<table>
<thead>
<tr>
<th>Measure</th>
<th>N</th>
<th>Correct Responses</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SPQ</strong></td>
<td></td>
<td>( \leq \frac{3}{4} ) = ( \frac{4}{4} )</td>
</tr>
<tr>
<td>Pre-training</td>
<td>294</td>
<td>30.30%</td>
</tr>
<tr>
<td>Post-training</td>
<td>304</td>
<td>16.80%</td>
</tr>
<tr>
<td>Follow-up</td>
<td>177</td>
<td>11.90%</td>
</tr>
<tr>
<td><strong>IKT-R</strong></td>
<td></td>
<td>( \leq \frac{7}{17} ) &gt; ( \frac{7}{17} )</td>
</tr>
<tr>
<td>Pre-training</td>
<td>124</td>
<td>52.40%</td>
</tr>
<tr>
<td>Post-training</td>
<td>137</td>
<td>16.10%</td>
</tr>
<tr>
<td>Follow-up</td>
<td>104</td>
<td>26.90%</td>
</tr>
<tr>
<td><strong>SOQ-R</strong></td>
<td></td>
<td>( \leq \frac{11}{20} ) &gt; ( \frac{11}{20} )</td>
</tr>
<tr>
<td>Pre-training</td>
<td>306</td>
<td>57.50%</td>
</tr>
<tr>
<td>Post-training</td>
<td>308</td>
<td>42.20%</td>
</tr>
<tr>
<td>Follow-up</td>
<td>178</td>
<td>42.10%</td>
</tr>
<tr>
<td><strong>SIRI-R</strong></td>
<td></td>
<td>( \leq \frac{6}{12} ) &gt; ( \frac{6}{12} )</td>
</tr>
<tr>
<td>Pre-training</td>
<td>116</td>
<td>50.90%</td>
</tr>
<tr>
<td>Post-training</td>
<td>134</td>
<td>29.10%</td>
</tr>
<tr>
<td>Follow-up</td>
<td>105</td>
<td>20.00%</td>
</tr>
</tbody>
</table>

\(^a\) = Comparison of pre-post performed with GEE \( p < .001 \)

\(^b\) = Comparison of pre-follow-up performed with GEE \( p < .001 \)

\(^c\) = Comparison of pre-follow-up performed with GEE \( p < .002 \)
Table 13: Immediate and Long-term Effects of the Training on Estimated Marginal Mean Scores

<table>
<thead>
<tr>
<th>Measure</th>
<th>Estimated Mean (lower bound)</th>
<th>SE</th>
<th>F</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPQ</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-training</td>
<td>3.56</td>
<td>0.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post training</td>
<td>3.84</td>
<td>0.04</td>
<td>&lt; .0001(^a)</td>
<td></td>
</tr>
<tr>
<td>Follow-up</td>
<td>3.85</td>
<td>0.04</td>
<td>17.135(^c)</td>
<td>&lt; .0001(^b)</td>
</tr>
<tr>
<td>IKT-R</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-training</td>
<td>7.18</td>
<td>0.19</td>
<td></td>
<td>&lt; .0001(^a)</td>
</tr>
<tr>
<td>Post-training</td>
<td>9.58</td>
<td>0.19</td>
<td></td>
<td>&lt; .0001(^a)</td>
</tr>
<tr>
<td>Follow-up</td>
<td>8.84</td>
<td>0.25</td>
<td>64.720(^c)</td>
<td>&lt; .0001(^b)</td>
</tr>
<tr>
<td>SOQ-R</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-training</td>
<td>10.31</td>
<td>0.19</td>
<td></td>
<td>&lt; .0001(^a)</td>
</tr>
<tr>
<td>Post-training</td>
<td>11.50</td>
<td>0.20</td>
<td></td>
<td>&lt; .0001(^a)</td>
</tr>
<tr>
<td>Follow-up</td>
<td>11.42</td>
<td>0.24</td>
<td>55.947(^c)</td>
<td>&lt; .0001(^b)</td>
</tr>
<tr>
<td>SIRI-R</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-training</td>
<td>6.29</td>
<td>0.23</td>
<td></td>
<td>&lt; .0001(^a)</td>
</tr>
<tr>
<td>Post-training</td>
<td>7.76</td>
<td>0.23</td>
<td></td>
<td>&lt; .0001(^a)</td>
</tr>
<tr>
<td>Follow-up</td>
<td>8.29</td>
<td>0.27</td>
<td>39.941(^c)</td>
<td>&lt; .0001(^b)</td>
</tr>
</tbody>
</table>

\(^a\) = Based on pre-post pairwise Sidak test comparisons
\
\(^b\) = Based on pre-follow-up pairwise Sidak test comparison
\
\(^c\) = Fixed effects of training (CS)
Table 14: Immediate and Long-term Effects of Training on Raw Mean Scores of Outcome Measures

<table>
<thead>
<tr>
<th>Measure</th>
<th>Pre-training</th>
<th>Post-training</th>
<th>Follow-up</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>P value(^c)</th>
<th>Effect Size(^d)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPQ</td>
<td>293</td>
<td>3.61</td>
<td>0.66</td>
<td>305</td>
<td>3.78</td>
<td>0.54</td>
<td>&lt; 0.001(^a)</td>
<td>0.41(^a)</td>
</tr>
<tr>
<td></td>
<td>177</td>
<td>3.87</td>
<td>0.37</td>
<td>177</td>
<td>3.87</td>
<td>0.37</td>
<td>&lt; 0.001(^b)</td>
<td>0.48(^b)</td>
</tr>
<tr>
<td>IKT-R</td>
<td>125</td>
<td>6.90</td>
<td>2.3</td>
<td>137</td>
<td>9.40</td>
<td>2.0</td>
<td>&lt; 0.001(^a)</td>
<td>1.25(^a)</td>
</tr>
<tr>
<td></td>
<td>106</td>
<td>8.60</td>
<td>2.0</td>
<td>106</td>
<td>8.60</td>
<td>2.0</td>
<td>&lt; 0.001(^b)</td>
<td>0.96(^b)</td>
</tr>
<tr>
<td>SOQ-R</td>
<td>305</td>
<td>10.93</td>
<td>2.9</td>
<td>307</td>
<td>12.34</td>
<td>3.2</td>
<td>&lt; 0.001(^a)</td>
<td>0.70(^a)</td>
</tr>
<tr>
<td></td>
<td>178</td>
<td>12.01</td>
<td>2.7</td>
<td>178</td>
<td>12.01</td>
<td>2.7</td>
<td>&lt; 0.001(^b)</td>
<td>0.60(^b)</td>
</tr>
<tr>
<td>SIRI-R</td>
<td>115</td>
<td>6.40</td>
<td>2.5</td>
<td>135</td>
<td>7.80</td>
<td>2.2</td>
<td>&lt; 0.001(^a)</td>
<td>0.89(^a)</td>
</tr>
<tr>
<td></td>
<td>105</td>
<td>8.60</td>
<td>2.2</td>
<td>105</td>
<td>8.60</td>
<td>2.2</td>
<td>&lt; 0.001(^b)</td>
<td>1.37(^b)</td>
</tr>
</tbody>
</table>

\(^a\) = Pre-post comparisons
\(^b\) = Pre-follow-up comparisons
\(^c\) = Based on post hoc pairwise comparisons using Sidak procedures
\(^d\) = Based on Cohen’s d for repeated measures
Table 15: Comparison of the Proportions of SPQ Correct Responses of the Occupational Groups

<table>
<thead>
<tr>
<th>Occupational Group</th>
<th>N</th>
<th>SPQ Correct Responses</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>≤ 4/4</td>
<td>=4/4</td>
<td></td>
</tr>
<tr>
<td>Special Constables</td>
<td>Pre-training</td>
<td>21</td>
<td>47.60%</td>
<td>52.40%</td>
</tr>
<tr>
<td></td>
<td>Post-training</td>
<td>22</td>
<td>0%</td>
<td>100.00%a</td>
</tr>
<tr>
<td></td>
<td>Follow-up</td>
<td>22</td>
<td>9.10%</td>
<td>90.90%b</td>
</tr>
<tr>
<td>Supervisors</td>
<td>Pre-training</td>
<td>58</td>
<td>27.60%</td>
<td>72.40%</td>
</tr>
<tr>
<td></td>
<td>Post-training</td>
<td>54</td>
<td>3.70%</td>
<td>96.30%c</td>
</tr>
<tr>
<td></td>
<td>Follow-up</td>
<td>31</td>
<td>19.40%</td>
<td>80.60%d</td>
</tr>
<tr>
<td>Train Operators</td>
<td>Pre-training</td>
<td>169</td>
<td>26.90%</td>
<td>73.10%</td>
</tr>
<tr>
<td></td>
<td>Post-training</td>
<td>160</td>
<td>21.20%</td>
<td>78.80%e</td>
</tr>
<tr>
<td></td>
<td>Follow-up</td>
<td>67</td>
<td>10.40%</td>
<td>89.60%f</td>
</tr>
</tbody>
</table>

\[ \text{a} = \text{Comparison pre-post Sidak procedure } p < .001 \]
\[ \text{b} = \text{Comparison pre-follow-up Sidak procedure } p < .001 \]
\[ \text{c} = \text{Comparison pre-post Sidak procedure } p < .001 \]
\[ \text{d} = \text{Comparison pre-follow-up Sidak procedure } p = .28 \]
\[ \text{e} = \text{Comparison pre-post Sidak procedure } p = .076 \]
\[ \text{f} = \text{Comparison pre-follow-up Sidak procedure } p = .002 \]

Based on the GEE, using unstructured working correlation matrix, the increase in proportion of Special Constables who correctly answered all four questions post-training and at follow-up was statistically significant (\( p < .001 \)). Correspondingly, as illustrated in Table 16, there was a statistically significant increase in the post-training SPQ mean score when compared to the pre-training mean score (\( \text{M} = 3.48 \) vs. \( \text{M} = 4.0 \), \( p < .001 \)) and the Cohen’s d test for repeated measure indicated a large effect size corresponding to 1.2. The absolute change in the
mean score (pre-follow-up) was +.52 and the observed relative change in the mean score was 14.9%. Likewise, at the 3-month follow-up there was a statistically significant increase in the mean score ($M = 3.48$ vs. $M = 3.91$, $p = .001$) and the large effect size was maintained (Cohen’s $d = .91$). The absolute change in the mean score (pre-follow-up) was +.43, with an observed relative increase in the SPQ follow-up mean score of 12.4%.

Similarly, an increase in the number of Transportation Supervisors who correctly answered all four SPQ questions is noted post-training and at the 3-month follow-up. The proportion of Transportation Supervisors with four correct responses increased post-training from 72.4% to 96.3%. At the 3-month follow-up, 80.6% of the Transportation Supervisors correctly answered correctly all four SPQ questions. Based on the GEE, using an unstructured working correlation matrix, the increase in post-training proportion of Transportation Supervisors who correctly answered all four questions was statistically significant ($p < .001$). No statistical significance was observed at the three-month follow-up ($p = .28$). Correspondingly, there was a statistically significant increase in the post-training SPQ mean score from the pre-training mean score ($M = 3.66$ vs. $M = 3.96$, $p = .003$), and the Cohen’s $d$ test for repeated measure indicated a medium training effect corresponding to .62. The observed absolute change in the mean score was +.30 with a relative mean score change of 8.2%.

There was a very slight increase in mean score at the three-month follow-up ($M = 3.65$ vs. $M = 3.78$); however, the increase was not statistically significant ($p = .62$) and the effect of the training was minimal (Cohen’s $d = .09$). The absolute change in the mean score (pre- vs. follow-up) was +.05 and the observed relative mean score change was 3%.

As illustrated in Table 15, the proportion of Train Operators who correctly answered all four SPQ questions increased slightly post-training from 73.1% to 78.8%; however, based on the GEE the increase observed was not statistically significant ($p = .076$). Conversely, a statistically significant increase in the number of Train Operators with four correct SPQ answers was noted at the three-month follow-up (73.1 to 89.6% $p = .002$; $M = 3.66$ vs. $M = 3.73$, $p = .67$), and the Cohen’s $d$ test for repeated measures indicated a small effect size corresponding to .29. The absolute change observed in the mean score was +.07 with a relative change in the mean score of 1.9%.
Table 16: Immediate and Long-term Effects of Training on the SPQ Raw Mean Scores of the Occupational Groups

<table>
<thead>
<tr>
<th>Occupational Group</th>
<th>SPQ Data</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Mean</td>
<td>SD</td>
<td>P value&lt;sup&gt;c&lt;/sup&gt;</td>
<td>Effect Size&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>Special Constables</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-training</td>
<td>21</td>
<td>3.48</td>
<td>0.60</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-training</td>
<td>23</td>
<td>4.00</td>
<td>0.00</td>
<td>&lt;0.001&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1.23&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Follow-up</td>
<td>22</td>
<td>3.91</td>
<td>0.29</td>
<td>0.001&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.91&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Transportation Supervisors</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-training</td>
<td>57</td>
<td>3.66</td>
<td>0.64</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-training</td>
<td>56</td>
<td>3.96</td>
<td>0.19</td>
<td>0.003&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.62&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Follow-up</td>
<td>32</td>
<td>3.77</td>
<td>0.50</td>
<td>0.61&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.09&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Train Operators</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-training</td>
<td>157</td>
<td>3.66</td>
<td>0.62</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-training</td>
<td>166</td>
<td>3.73</td>
<td>0.58</td>
<td>0.67&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.29&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Follow-up</td>
<td>72</td>
<td>3.89</td>
<td>0.36</td>
<td>0.024&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.70&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

<sup>a</sup> = Pre-post comparisons  
<sup>b</sup> = Pre-follow-up comparisons  
<sup>c</sup> = Based on Post Hoc pairwise comparisons using Sidak procedures  
<sup>d</sup> = Based on Cohen’s d for repeated measures

As indicated in Table 16, there was a slight non-statistically significant increase in the post-training SPQ mean score when compared to the pre-training mean score ($M = 3.66$ vs. $M =$ The comparison of pre vs. follow-up mean scores indicated a statistically significant increase in
the SPQ mean at the three-month follow-up ($M = 3.66$ vs. $M = 3.89$, $p = .009$) with a moderate effect size (Cohen’s $d = .70$). The observed absolute change in the mean score was $+.23$, with a relative increase in the SPQ mean score at the three-month follow-up of $6.3\%$.

To assess the impact of occupation i.e.: Special Constables, Transportation Supervisors, and Train Operators on the SPQ (outcome measure), a secondary repeated measures model was performed adding a term for Time by Occupation (independent variables). The linear mixed-effects analysis for was performed using CS for correlation structure. Contrast tests permitting post-hoc pair-wise comparisons between Time and Occupational group means were performed using a Sidak approach.

The analysis of variance showed that the interaction between the training and occupational group (Time X Occupation) was significant: $F_{6, 497} = 3.682$, $p = 0.001$. Post hoc contrast analyses showed no statistically significant differences in the estimated marginal means between the Train Operators, Special Constables and Transportation Supervisor at pre-training ($p = .698$ and $p = 1.0$); however, post–training Train Operators had lower estimated marginal mean scores differences compared to those of Special Constables ($p = .077$) and lower than Transportation Supervisors ($p = .064$), albeit statistically non-significant. The differences in the estimated marginal mean scores at the three–month follow-up were not statistically significant (Train Operators vs. Special Constables $p = 1.0$; Train Operators vs. Transportation Supervisors $p = .948$; Special Constables vs. Transportation Supervisors $p = .950$).

As demonstrated in Table 17, the greatest difference in pre-post estimated marginal mean scores was observed in the pre-post scores of Special Constables ($\Delta = .54$, $p = .001$) with the greatest effect ($\Delta/SE$ of 3.86. The difference and effect observed for Transportation Supervisors was slightly lower ($\Delta = .30$, $p = .002$, $\Delta/SE = 3.39$). Conversely, for Train Operators the difference in post-pre estimated marginal mean scores was not statistically significant ($p = .43$). At the three-month follow-up Special Constables continued to lead with the greatest difference in follow-up-pre estimated marginal mean scores ($\Delta = .41$, $p = .021$) and effect ($\Delta/SE = 2.73$), followed by Train Operators ($\Delta = .21$, $p = .009$, $\Delta/SE = 2.93$). The difference for Transportation Supervisors was not statistically significant ($p = .62$). Figure 2 illustrates the interaction between occupational group affiliation and the immediate and long-term effects of the training on SPQ mean scores.
The training had a differential effect on the suicide-risk procedural knowledge acquisition of the three occupational groups. After the training, there were significant differences in the magnitude of the gains in knowledge made by Special Constables, Transportation Supervisors, and Train Operators. These differences are demonstrated in the variations observed in the occupational groups’ mean scores and the corresponding effect sizes, the magnitude of the relative increases in mean scores.

Figure 2: The interaction between occupational group affiliation and the immediate and long-term effects of the training on SPQ mean scores

![Graph showing the interaction between occupational group affiliation and the immediate and long-term effects of the training on SPQ mean scores.]

**Intervention Knowledge Test -R (IKT-R)**

One of the study’s objectives was to determine whether participants’ intervention knowledge increased as a result of attending the suicide prevention workshops. The IKT-R questionnaire was only applicable to those participants who attended the *safeTALK* workshop; thus, the questionnaire was not included in the study packets distributed to *suicideAWARE* attendees. The IKT-R measured salient aspects of three of the major modules of the *SafeTALK* training: attitudes, knowledge, and intervention skills. The modified IKT version used in the study consisted of 17 multiple choice questions. The case study questions which were included
in the original IKT were eliminated as were questions directly related to the ASIST workshop material (Tierney, 1988). Scores on the Intervention Knowledge Test Questionnaire-Revised (IKT-R) were computed by combining the correct responses with a maximum possible total score of 17. For the purposes of the Generalized Estimation Equation analysis two categories were created using the pre-training median score of 7: “≤ 7” and “>7 correct answers” with the former coded as “0” and the latter coded as ”1”. A major question in this study related to whether or not participants’ knowledge regarding suicide intervention changed as a result of the training experience. As illustrated on Table 12, post-training the proportion of study participants who correctly answered seven or more of the 17 questions on the IKT-R questionnaire increased from 47.6% at pre-training to 83.9% and at the three--month follow-up the proportion of participants with seven or more correct answers was greater than observed at pre-training (73.1% vs. 47.6%).

To compare the changes in proportions of study participants with correct answers, the GEE was performed using an unstructured working correlation matrix with post-hoc Sidak procedures. The changes in proportions pre-post and pre-follow-up were statistically significant (p < .001 and p = .002, respectively). Comparing the characteristics of participants who had low IKT-R (≤ 7/17 ) scores (n = 64) with those who had high IKT-R (> 7/17) scores (n = 60) indicated the two groups differed moderately in terms of demographic characteristics, such as age (> 35; 86% vs. 80%, respectively), gender (male: 86% vs. 74%, respectively), and occupational group membership with a larger proportion of those with lower IKT-R scores employed as Transportation Supervisors. The two groups did not differ in terms of years of employment at the TTC (< 10 years 84% vs. 82%, respectively), experience with MHA apprehensions (0 experience 68% vs. 68%, respectively), or experience with suicide and suicidal behaviour (0 experience: 59% vs. 54%, respectively).

To examine the effect of the training on acquisition of intervention knowledge a mixed linear fixed-effects analysis for repeated measures was performed using a CS correlation matrix. The analysis of variance showed that the effects of the training on mean scores were significant (F2,245 = 64.720, p < .0001( see Table 13). Further post hoc analyses using Sidak pairwise comparisons procedure adjusting for multiple testing indicated statistically significant changes in the estimated marginal mean score of IKT-R post-training with the pre-training mean score significantly lower than the mean score observed immediately post training (M = 7.18, SE .19
As illustrated in Table 14, further comparisons of pre- and post-training mean scores using Cohen’s d for repeated measures indicated the training had a large effect on the IKT-R post-training mean score (Cohen’s d = 1.25) with an absolute change in pre-post-training mean score (M = 6.9 vs. M = 9.4, p < .001) of +2.5 and a relative increase in the post-training mean score of 36.2%. Similarly, comparisons of pre vs. follow-up mean scores indicated the training continued to have a large effect on IKT-R scores (Cohen’s d = .96) with an absolute change in

Table 17: SPQ Model Based on Differences

<table>
<thead>
<tr>
<th>Occupational Group</th>
<th>Difference</th>
<th>SE</th>
<th>Effect (D/SE)</th>
<th>DF</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Special Constables</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-Pre</td>
<td>0.54</td>
<td>0.14</td>
<td>3.86</td>
<td>451</td>
<td>0.001</td>
</tr>
<tr>
<td>FU-Pre</td>
<td>0.41</td>
<td>0.15</td>
<td>2.73</td>
<td>578</td>
<td>0.021</td>
</tr>
<tr>
<td>FU-Post</td>
<td>-0.14</td>
<td>0.15</td>
<td>-0.94</td>
<td>577</td>
<td>0.72</td>
</tr>
<tr>
<td>Transportation Supervisors</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-Pre</td>
<td>0.30</td>
<td>0.09</td>
<td>3.39</td>
<td>422</td>
<td>0.002</td>
</tr>
<tr>
<td>FU-Pre</td>
<td>0.12</td>
<td>0.11</td>
<td>1.09</td>
<td>521</td>
<td>0.62</td>
</tr>
<tr>
<td>FU-Post</td>
<td>-0.18</td>
<td>0.11</td>
<td>-1.64</td>
<td>540</td>
<td>0.26</td>
</tr>
<tr>
<td>Train Operators</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-Pre</td>
<td>0.07</td>
<td>0.05</td>
<td>0.14</td>
<td>421</td>
<td>0.43</td>
</tr>
<tr>
<td>FU-Pre</td>
<td>0.21</td>
<td>0.07</td>
<td>2.93</td>
<td>514</td>
<td>0.009</td>
</tr>
<tr>
<td>FU-Post</td>
<td>0.14</td>
<td>0.07</td>
<td>2.00</td>
<td>505</td>
<td>0.13</td>
</tr>
</tbody>
</table>
pre vs. follow-up mean score ($M = 6.9$ vs. $M = 8.6$, $p < .001$) of +1.7 with a relative increase in the follow-up mean score of 24.6%.

**Linked vs. Unlinked Scores**

Comparisons of the linked and unlinked IKT-R data using one-way ANOVA tests indicated a statistical difference between the pre-training mean scores of the linked and unlinked IKT-R scores ($p = .004$); however, further comparison of post-training and three-month follow-up linked and unlinked mean scores indicated the differences were not statistically significant ($p = .096$, $p = .055$, respectively). As further illustrated on Table A, Appendix N, the additional mixed linear fixed-effect analyses indicate there were no differences in the effect of the training on linked and unlinked data.

**IKT-R: Occupational Groups**

As illustrated in Table 18, prior to the training slightly more than half of the Special Constables (55%) had correctly answered more than 7 questions on the Intervention IKT-R questionnaire whereas post-training the proportion of Special Constables with more than seven correct answers increased to 81.8%. The long term effects of the training were also apparent at the three-month follow-up with more the three-quarter of the Special Constables (76.2%) correctly answered more than 7 IKT-R questions. Based on the GEE, using an unstructured working correlation matrix, the increase in proportions of Special Constables who post-training answered correctly more than seven questions was statistically significant ($p = .037$); however, the differences in proportions pre vs. follow-up were not ($p = .142$). Correspondingly, as illustrated in Table 19, there was a statistically insignificant increase in the Special Constables post-training IKT-R mean score ($M = 7.4$ vs. $M = 9.2$, $p = .045$). The Cohen’s $d$ test for repeated measures indicated a large effect size corresponding to .92. The absolute change in the IKT-R post-training mean score was +1.8, and the relative increase in the post-training mean score was 16.8%. Although the raw mean score at the three-month follow-up indicated an increase from the pre-training score ($M = 7.4$ vs. $M = 8.7$, $p = .24$) the difference was not statistically significant; however, the large effect size was maintained (Cohen’s $d = .80$). The absolute change in the IKT-R follow-up mean score was +1.3 and the relative increase was 17.6%. Post-training an increase in the number of Transportation Supervisors who answered correctly seven
or more questions is noted. The proportion of Transportation Supervisors with more than seven correct responses increased from 34.6% pre-training to 81.5% post-training. The effects of the training were maintained at the three-month follow-up with nearly two-thirds (64.5%) of the Transportation Supervisors who correctly answered more than 7 IKT-R questions. Based on the GEE, using an unstructured working correlation matrix, the increase in proportion of Transportation Supervisors who correctly answered seven or more IKT-R questions post-training and at follow-up was statistically significant (p = .001 and p = .002, respectively).

Table 18: Comparison of Proportions of IKT-R Correct Responses of the Occupational Groups

<table>
<thead>
<tr>
<th>Occupational Group</th>
<th>N</th>
<th>≤ 7/17 Correct Responses</th>
<th>&gt; 7/17 Correct Responses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Special Constables</td>
<td>20</td>
<td>45.00%</td>
<td>55.00%</td>
</tr>
<tr>
<td>Pre-training</td>
<td>22</td>
<td>18.20%</td>
<td>81.80%&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Post-training</td>
<td>21</td>
<td>23.80%</td>
<td>76.20%&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Follow-up</td>
<td>21</td>
<td>23.80%</td>
<td>76.20%&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Transportation Supervisors</td>
<td>52</td>
<td>65.4%</td>
<td>34.60%</td>
</tr>
<tr>
<td>Pre-training</td>
<td>54</td>
<td>18.50%</td>
<td>81.50%&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>Post-training</td>
<td>31</td>
<td>35.50%</td>
<td>64.50%&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

<sup>a</sup> = Comparison pre-post Sidak procedure p = .037  
<sup>b</sup> = Comparison pre-follow-up Sidak procedure p = .142  
<sup>c</sup> = Comparison pre-post Sidak procedure p < .001  
<sup>d</sup> = Comparison pre-follow-up Sidak procedure p < .002

As illustrated on Table 19, compared to the pre-training mean score a statistically significant increase in the post-training IKT-R mean score is observed (M = 6.3 vs. M = 9.0, p
The Cohen’s d test for repeated measures indicated the training had large effect \( (d = 1.4) \). The absolute change in the post-training IKT-R mean score was +2.9, and the relative increase in the post-training mean score was 42.9%. Compared to the pre-training mean score a statistically significant increase in the three-month follow-up IKT-R mean score is observed \( (M = 6.3 \text{ vs. } M = 8.1, p = .001) \). The Cohen’s d test for repeated measures indicated the training had a large effect at the three-month follow-up \( (d = .96) \). The absolute change in the follow-up IKT-R mean score was +1.8 and the relative increase in the follow-up mean score was 28.6%. To establish whether there were statistically significant differences between the linked and unlinked mean scores of the different occupational groups’ one-way ANOVA tests were conducted. The comparisons indicated only a statistically significant differences in the pre-training linked and unlinked IKT-R scores of Special Constables \( (p = .027) \). No statistically significant differences were observed between the Transportation Supervisors’ linked and unlinked data.

To assess the impact of occupation on the IKT-R (outcome measure), secondary repeated measures models were performed adding a term for Time by Occupation (independent variables). Contrast tests analyses permitting post-hoc pairwise comparisons between time and occupation means were also performed using Sidak procedures.

The analysis of variance showed that the interaction between the training and occupational group affiliation \( (Time \times Occupation) \) was not significant: \( F_{4,235} = .344, p = .848 \). Table 20 presents the IKT-R model based on differences. As demonstrated in Table 20, post-hoc analyses revealed the greatest difference in post-pre estimated marginal mean scores was observed in the pre-post scores of the Transportation Supervisors \( (\Delta = 2.62, p = .001) \) with the greatest effect \( (\Delta/SE) \) of 7.94. The difference in pre-post estimated marginal mean scores for the Special Constables was statistically significant \( (\Delta = 1.90, p = .001, \Delta/SE = 3.58) \). At the three-month follow-up Transportation Supervisors continued to lead with the greatest difference in follow-up vs. pre-training estimated marginal mean scores \( (\Delta = 1.61, p = .001) \) and effect \( (\Delta/SE = 3.93) \), followed by Special Constables \( (\Delta = 1.33, p = .064, \Delta/SE = 2.29) \).

The second study hypothesis predicted positive attitudes towards suicide and suicide intervention would increase after the training and would be maintained over time. To investigate the second study hypothesis, attitudes were assessed using a modified version of the Suicide Opinion Questionnaire (SOQ) \( (\text{Domino et al, 1982}) \). The original version of the SOQ consists of 100 items that cover a wide range of attitudes towards suicide. For this study several revisions
were made to the SOQ in order to streamline the measure and to make it relevant for use with TTC employees. The primary revision to the SOQ involved streamlining the SOQ and selecting only 20 items from the standard 100 items. The revised SOQ-R is presented in Appendix D & E. Each item on the SOQ-R was rated on a five-point Likert scale from “Strongly agree” to “Strongly disagree.”

Table 19: Comparison by Occupational Group of the Effects of Training on IKT-R Raw Mean Scores

<table>
<thead>
<tr>
<th>Occupational Group</th>
<th>IKT-R Data</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>P Valuec</th>
<th>Effect Size(^d)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Special Constables</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-training</td>
<td>20</td>
<td>7.4</td>
<td>2.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-training</td>
<td>22</td>
<td>9.2</td>
<td>2.6</td>
<td>.045(^a)</td>
<td>0.92(^a)</td>
<td></td>
</tr>
<tr>
<td>Follow-up</td>
<td>22</td>
<td>8.7</td>
<td>2.2</td>
<td>.24(^b)</td>
<td>0.80(^b)</td>
<td></td>
</tr>
<tr>
<td>Transportation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Supervisors</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-training</td>
<td>53</td>
<td>6.3</td>
<td>2.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-training</td>
<td>54</td>
<td>9</td>
<td>2</td>
<td>&lt; .001(^a)</td>
<td>1.39(^a)</td>
<td></td>
</tr>
<tr>
<td>Follow-up</td>
<td>31</td>
<td>8.1</td>
<td>2.1</td>
<td>&lt; .001(^b)</td>
<td>0.96(^b)</td>
<td></td>
</tr>
</tbody>
</table>

\(^a\) = Pre-post comparisons  
\(^b\) = Pre-follow-up comparisons  
\(^c\) = Based on post hoc pairwise comparisons using Sidak procedures  
\(^d\) = Based on Cohen’s d for repeated measures
Table 20: IKT-R Model Based on Differences

<table>
<thead>
<tr>
<th>Occupational Group</th>
<th>Difference</th>
<th>SE</th>
<th>Effect (D/SE)</th>
<th>DF</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Special Constables</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-Pre</td>
<td>1.90</td>
<td>0.53</td>
<td>3.58</td>
<td>198</td>
<td>0.001</td>
</tr>
<tr>
<td>FU-Pre</td>
<td>1.33</td>
<td>0.58</td>
<td>2.29</td>
<td>285</td>
<td>0.064</td>
</tr>
<tr>
<td>FU-Post</td>
<td>-0.58</td>
<td>0.55</td>
<td>-1.05</td>
<td>279</td>
<td>0.65</td>
</tr>
<tr>
<td>Transportation Supervisors</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-Pre</td>
<td>2.62</td>
<td>0.33</td>
<td>7.94</td>
<td>186</td>
<td>0.001</td>
</tr>
<tr>
<td>FU-Pre</td>
<td>1.61</td>
<td>0.41</td>
<td>3.93</td>
<td>257</td>
<td>0.001</td>
</tr>
<tr>
<td>FU-Post</td>
<td>-1.01</td>
<td>0.40</td>
<td>-2.52</td>
<td>258</td>
<td>0.037</td>
</tr>
</tbody>
</table>

For the purpose of analysis the categories were collapsed into three categories by combining the “Strongly Agree” and “Agree” categories, the “Strongly disagree” and “Disagree” categories, and “Undecided”. The participant's total score consisted of the number of positive attitudes he/she endorsed. For example, if the participant endorsed a statement that was indicative of negative attitude (e.g., Question #3: “Those who threaten suicide rarely do.”) by agreeing with the statement he/she scored 0, while if they disagreed with the statement he/she scored 1. Similarly, if the participant endorsed a positive attitude (e.g., Question #19: “Potentially, everyone can be a suicide victim.”) by agreeing with the statement he/she scored 1, while disagreeing they scored 0. The maximum possible score for the SOQ-R was 20.

For the purposes of the Generalized Estimation Equation analysis two categories were created “≤ 11 positive attitudes” and “> 11 positive attitudes” with the former coded as “0” and the latter coded as a ”1”. The categories were created using the pre-training median score of 11. As indicated in Table 12, prior to the suicide prevention training less than half (42.5%) of the participants endorsed a greater number of positive attitudes (> 11) towards suicide prevention. Post-training the proportion of employees with a greater number of positive attitudes increased
to 57.8%. The increase in the proportion of employees endorsing a greater number of positive attitudes was maintained at the three-month follow-up with 57.9% of the participants endorsing more than 11 positive attitudes. Using an unstructured correlation matrix in the GEE analysis, the post-hoc analyses using Sidak procedure indicated the pairwise comparisons of estimated marginal means (pre-post and pre-follow-up) were statistically significant (p < .001).

Analyzing the characteristics of participants who had low SOQ (≤11/20) scores (n = 174) with those who had high SOQ (>11/20) scores (n = 132) suggested the two groups did not differ in demographic characteristics such as age (> 35; 82% vs. 87%, respectively), gender (male: 85% vs. 89%, respectively), and years of employment at the TTC (< 10 years 82% vs. 81%), experience with MHA apprehensions (0 experience: 66% vs. 59%, respectively); however, a larger proportion of those with higher scores were Train Operators (30% vs. 72%, respectively), and they had no experience with suicide and suicidal behaviour (0 experience: 54% vs. 75%, respectively).

To examine the effect of the training on changes in attitudes towards suicide and the suicidal person, a mixed linear fixed-effects analysis for repeated measures was performed using a CS correlation matrix. The analysis of variance showed that the effects of the training on estimated marginal mean scores were significant (F2, 479 = 55.947, p < .001; Table 13). Further post-hoc analyses using Sidak pairwise comparisons procedure adjusting for multiple testing indicated statistically significant changes in the estimated marginal mean score of SOQ-R with the pre-training mean score significantly lower than the mean score observed immediately post-training (M = 10.31, SE = .19 vs. M = 11.5, SE = .20) and at three-month follow-up (M = 11.4, SE = .24) with an overall alpha level of < 0.001.

Further comparisons of pre- and post-training raw mean scores (see Table 14) using Cohen’s d for repeated measures indicated the training had a moderate effect on the SOQ-R post-training raw mean score (Cohen’s d = .7) with an absolute change in pre-post-training mean score (M = 10.93 vs. M = 12.34, p < .001) of +1.41 in the post-training mean score and a relative increase in the post-training mean score of 12.9%. Similarly, comparisons of pre vs. follow-up mean scores indicated the training continued to have a moderate effect (Cohen’s d = .6) with an absolute change in pre vs. follow-up mean score (M = 10.93 vs. M = 12.01, p < .001) of +1.08 and a relative increase in the follow-up score of 9.9%.
Linked vs. Unlinked Scores

Comparisons of SOQ-R linked and unlinked mean scores (See Table B in Appendix N) conducted with one-way ANOVA tests, indicated there were no statistically significant differences between the linked and unlinked SOQ-R mean scores ($F_{1,788} = 1.021$, $p = .313$). Table A in Appendix N presents the results of mixed linear fixed-effects analyses and the post-hoc pairwise comparisons of the linked and unlinked scores.

SOQ-R: Occupational Groups

As illustrated in Table 21, there was slight increase post-training in the number of Special Constables who answered correctly more than 11 of the SOQ-R questions (22.7% vs. 27.3%). This number further increased at the three-month follow-up to 36.4%; however, based on the GEE these increases in proportion were not statistically significant ($p = .74$, and $p = .312$, respectively). Correspondingly, as shown in Table 22, there was no statistically significant observed increase in the post-training mean score when compared to the pre-training mean score ($M = 9.86$ vs. $M = 10.36$, $p = .145$) with a small effect size (Cohen’s $d = .32$). The absolute change in the mean score was $.5$, and the relative increase in the post-training SOQ-R mean score was 5.1%. At the three-month follow-up the increase was not statistically significant although the mean score increased from $M = 9.86$ to $M = 10.77$ ($p = .241$) with a large effect size ($d = .74$). The absolute change in the mean score at the three-month follow-up was $+1.39$ with a relative increase in the SOQ mean score of 9.2%. The number of Transportation Supervisors who correctly answered more than eleven of the 20 SOQ-R questions post-training increased slightly post-training (22.4% vs. 24.1%), an increase which was not statistically significant ($p = .793$) based on an unstructured correlation matrix in the GEE analysis; however, at the three-months follow-up the increase observed in the proportion of Transportation Supervisors with more than eleven correct SOQ-R answers (22.4% vs. 49.9%) was statistically significant ($p = .028$).

As indicated in Table 22, the increase in the post-training mean score compared to the pre-training mean score was not statistically significant ($M = 9.66$ vs. $M = 10.20$ $p = .454$), and the Cohen’s $d$ effect size was small ($d = .33$). The absolute change in the post-training mean score was $+.56$ and the relative increase in the SOQ-R mean score was 5.6%. Similarly, the increase in mean score continued at the three-month follow-up; however, it persisted to be non-
significant ($M = 9.66$ vs. $M = 10.81$, $p = .054$), and the Cohen’s $d$ repeated measures calculations indicated a medium effect size ($d = .64$). The absolute change in the SOQ-R mean score was +.91 with a relative increase in the mean score of 11.9%.

The number of Train Operators who correctly answered more than eleven of the 20 SOQ-R questions increased significantly from pre- to post-training (58.4% to 82.2% \( p < .001 \)), a trend that continued and was evident at the three-month follow-up (58.4% vs. 85.1%. \( p < .001 \)). Correspondingly, there was a statistically significant increase in the post-training SOQ-R mean score when compared to the pre-training mean score ($M = 12.52$ vs. $M = 14.0$, $p = .001$), and a large effect size ($d = .72$). The absolute change in the post-training mean score was +1.48 and the relative increase in the SOQ-R post training score was 11.8%. The comparison of pre- vs. follow-up mean scores indicated there was an increase in the SOQ-R mean at the three-months follow-up ($M = 12.52$ vs. $M = 13.76$, $p = .001$), the effect size was medium with a Cohen’s $d = .63$. The absolute change in the SOQ-R post-training mean score was +1.24 and the relative increase in the follow-up mean score was 9.9%.

Comparisons of the Special Constables and the Transportation Supervisors and Train Operators linked and unlinked SOQ-R mean scores indicated there was no statistically significant difference between the linked and unlinked mean scores ($p = .077$ and $p = .094$, $p = .713$, respectively).

To assess the impact of occupation i.e., Special Constables, Transportation Supervisors, and Train Operators on the SOQ-R (outcome measure), a secondary repeated measures model was performed adding a term for Time by Occupation (independent variables). The linear mixed-effects analysis was performed using the CS correlation structure. Contrast tests permitting post-hoc pair-wise comparisons between Time and Occupational group means were performed using Sidak procedures. The secondary analysis of variance indicated the interaction between the training and occupational group (Time X Occupation) was significant: ($F_{6,494} = 3.640$, $p = 0.002$). The post-hoc contrasts analyses showed statistically significant differences in the Train Operators estimated marginal mean scores compared to those of Special Constables and Transportation Supervisors (\( p < .001 \)); however, the differences between the estimated marginal mean scores of Special Constables and Transportation Supervisors were not statistically significant. Figure 3 illustrates the interaction between the occupational group affiliation and the immediate and long-terms effects of the training on SOQ-R mean scores.
Table 21: Comparison of the Proportions of SOQ-R Correct Responses of the Occupational Groups

<table>
<thead>
<tr>
<th>Occupational Group</th>
<th>N</th>
<th>SOQ-R Correct Responses</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>≤ 11/20</td>
<td>&gt; 11/20</td>
<td></td>
</tr>
<tr>
<td>Special Constables</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-training</td>
<td>22</td>
<td>77.30%</td>
<td>22.70%</td>
<td></td>
</tr>
<tr>
<td>Post-training</td>
<td>22</td>
<td>72.70%</td>
<td>27.30%a</td>
<td></td>
</tr>
<tr>
<td>Follow-up</td>
<td>22</td>
<td>63.60%</td>
<td>36.40%b</td>
<td></td>
</tr>
<tr>
<td>Transportation Supervisors</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-training</td>
<td>58</td>
<td>77.60%</td>
<td>22.40%</td>
<td></td>
</tr>
<tr>
<td>Post-training</td>
<td>54</td>
<td>75.90%</td>
<td>24.10%c</td>
<td></td>
</tr>
<tr>
<td>Follow-up</td>
<td>31</td>
<td>58.10%</td>
<td>41.90%d</td>
<td></td>
</tr>
<tr>
<td>Train Operators</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-training</td>
<td>166</td>
<td>41.60%</td>
<td>58.40%</td>
<td></td>
</tr>
<tr>
<td>Post-training</td>
<td>163</td>
<td>17.80%</td>
<td>82.20%e</td>
<td></td>
</tr>
<tr>
<td>Follow-up</td>
<td>67</td>
<td>14.90%</td>
<td>85.10%f</td>
<td></td>
</tr>
</tbody>
</table>

*a = Comparison pre-post Sidak procedure p = .74
*b = Comparison pre-follow-up Sidak procedure p = .312
*c = Comparison pre-post Sidak procedure p = .793
*d = Comparison pre-follow-up Sidak procedure p = .028
*e = Comparison pre-post Sidak procedure p < .001
*f = Comparison pre-follow-up Sidak procedure p < .001
Table 22: The Effects of the Training on the SOQ-R Raw Mean Scores of the Various Occupational Groups

<table>
<thead>
<tr>
<th>SOQ-R Data</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>P value&lt;sup&gt;c&lt;/sup&gt;</th>
<th>Effect Size&lt;sup&gt;d&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Occupational Group</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Special Constables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-training</td>
<td>22</td>
<td>9.86</td>
<td>2.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-training</td>
<td>23</td>
<td>10.36</td>
<td>2.3</td>
<td>.83&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.32&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Follow-up</td>
<td>22</td>
<td>10.77</td>
<td>1.9</td>
<td>.42&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.74&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td><strong>Transportation Supervisors</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-training</td>
<td>57</td>
<td>9.66</td>
<td>1.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-training</td>
<td>56</td>
<td>10.20</td>
<td>2.4</td>
<td>.454&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.33&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Follow-up</td>
<td>32</td>
<td>10.81</td>
<td>1.9</td>
<td>.054&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.64&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td><strong>Train Operators</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-training</td>
<td>167</td>
<td>12.52</td>
<td>3.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-training</td>
<td>169</td>
<td>14.00</td>
<td>2.7</td>
<td>&lt;.001&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.72&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Follow-up</td>
<td>73</td>
<td>13.76</td>
<td>2.8</td>
<td>&lt;.001&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.63&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

<sup>a</sup> = Pre-post comparisons  
<sup>b</sup> = Pre-follow-up comparisons  
<sup>c</sup> = Based on post hoc pairwise comparisons using Sidak procedures  
<sup>d</sup> = Based on Cohen’s d for repeated measures
Figure 3. The interaction between occupational group affiliation and immediate and long-term training effects on the SOQ-R mean scores.

Table 23 presents the SOQ-R Model based on differences. As demonstrated in Table 23, the greatest difference in post-pre estimated marginal mean scores among the three occupational groups was the difference observed in the pre-post scores of Train Operators ($\Delta = .20$, $p = .001$) with the greatest effect ($\Delta/SE$) of 11.0. Conversely, the difference in the estimated marginal mean scores and effect observed for Transportation Supervisors was much lower and was not statistically significant ($\Delta = .73$, $p = .10$, $\Delta/SE = 2.15$). Similarly, the difference in post-pre estimated marginal mean scores for the Special Constables was not statistically significant ($p = .67$). At the three-month follow-up Train Operators continued to lead with the greatest difference in follow-up-pre estimated marginal mean scores ($\Delta = 1.70$, $p = .001$) and effect ($\Delta/SE = 6.3$), followed by Transportation Supervisors ($\Delta = .95$, $p = .09$, $\Delta/SE = 2.21$). The difference for Special Constables was not statistically significant ($p = .36$).
The third hypothesis predicted participation in the *safeTALK* workshop would have positive immediate and long-term effects on participants’ suicide assessment and intervention skills. The modified Suicide Intervention Response Inventory (SIRI-R) was used to test the third study hypothesis.

The SIRI-R measured whether *safeTALK* participants’ ability to recognize facilitative and deleterious intervention responses in suicide intervention situations changed as a result of workshop attendance. The Suicide intervention Response Inventory consisted of six hypothetical statements that may be indicative of patron’s distress and suicide risk, and provides two possible helper responses for each patron’s statement (Helper A and Helper B), one of which demonstrates relatively more skillful management of the suicidal patron. Participants were asked
to identify which of the helpers’ statements were facilitative intervention responses and which statements were deleterious intervention responses. Facilitative responses were to be assigned a positive value (+1 to +3) whereas deleterious responses were to be assigned a negative value (-1 to -3). The scoring of the participants’ responses was based on the replies identified by the clinicians as facilitative or deleterious. The total score on the SIRI-R represents the number of correct responses the participant identified.

The method of coding and scoring were modified for the purpose of this study. It was felt that since the questionnaire was modified and the number of statements was substantially reduced from 25 to 6, the method of scoring suggested for the SIRI by Neimeyer and Bonnelle (1997) could not be applied. It was decided to revert to a scoring method suggested by Neimeyer and MacInnes in an earlier publication (1983). To be able to tally up the correct replies, participants’ responses were recorded prior to analysis. Participants’ replies that correctly identified the helper’s response were re-coded and scored as 1, whereas incorrect replies were re-coded as 0. A participant’s total score consisted of the number of responses correctly identified. The maximum possible score for the six statements was twelve.

As illustrated in Table 12, post-training the proportion of study participants who correctly identified six or more facilitative and deleterious intervention statements on the SIRI-R questionnaire increased from 49.1% at pre-training to 70.9%. At the three-month follow-up a further increase in proportion is observed (80.0%) among those participants with six or more correct responses. Using an exchangeable correlation matrix in the GEE analyses, the post-hoc Sidak pairwise comparisons procedures (pre- vs. post and pre- vs. follow-up) indicated the observed increase in proportion was statistically significant (p < .001).

Analyzing the characteristics of participants who had low SIRI-R (≤ 6/12) scores (n = 59) with those who had high SIRI-R (> 6/12) scores (n = 57) indicated the two groups differed in demographic characteristics, such as age (> 35; 95% vs. 75%, respectively, p < .001), gender (male: 86% vs. 73%, respectively, p = .07), years of employment at the TTC (< 10 years 86% vs. 80%, respectively, p = .2). A larger proportion of those with lower SIRI-R scores were employed as Transportation Supervisors (48% vs. 34%, respectively p = .025). They had no experience with suicide and suicidal behaviour (0 experience; 58% vs. 52%, respectively p = .96), and no experience with MHA apprehensions (0 experience; 71% vs. 66%, respectively, p = .52),
As indicated in Table 13, statistically significant effects of the training on the SIRI-R mean score is noted ($F_{2,213} = 39.927$, $p < .001$). Further post hoc analyses using Sidak pairwise comparisons procedures adjusting for multiple testing indicated statistically significant changes in the estimated marginal mean score with the pre-training score significantly lower than the mean score observed immediately post-training ($M = 6.29$, $ES = .23$ vs. $M = 7.76$, $ES = .23$) and at three-month follow-up ($M = 8.29$, $ES = .27$) with an overall alpha level $< .001$.

As illustrated in Table 14, further comparisons of pre-post-training scores using Cohen’s $d$ for repeated measures indicated the training had a large effect on the SIRI-R post-training mean score ($d = .89$) with an absolute change in post-training mean score of +1.40, and a relative increase in the post-training SIRI-R mean score of 21.9%. Likewise, comparisons of the pre- and follow-up mean scores showed the training continued to have a large effect (Cohen’s $d = 1.37$) with an absolute change in the three-month follow-up mean score of +2.2 and a relative increase in the follow-up score of 34.4%.

**Linked vs. Unlinked Scores**

Comparisons of the linked and unlinked data using one-way ANOVA tests indicated no statistical differences between the mean scores of the two datasets. Further post-hoc analyses using Sidak procedures comparing the linked and unlinked pre-, post, and follow-up mean scores indicated the differences were not statistically significant ($p = .820$, $p = .238$, $p = .071$, respectively). Further, comparisons of Special Constables and Transportation Supervisors linked and unlinked SIRI-R mean scores indicated there was no statistically significant difference between the mean scores ($p = .077$ and $p = .094$, respectively). Table A in appendix N presents the results of the mixed linear fixed-effects analyses and Sidak procedure results for the linked and unlinked data.

**SIRI-R: Occupational Groups**

As illustrated in Table 24, at post-training the proportion of Special Constables who correctly identified six or more facilitative and deleterious intervention statements on the SIRI-R questionnaire increased from the pre-training of 52.6% to 59.1%, and at the three-month follow-up the proportion increased to 72.7%. Based on the GEE, using an unstructured working correlation matrix, the observed increase in proportion was not statistically significant (pre vs.
The proportion of Transportation Supervisors who correctly identified six or more facilitative and deleterious intervention statements significantly increased post-training from 38.3% to 69.5% ($p < .001$) and at the three-month follow-up from 38.3% to 64.5% ($p = .009$).

Correspondingly, as illustrated in Table 25, there was a statistically insignificant increase in the Special Constables’ pre-training SIRI-R mean score immediately after the training ($M = 6.53$ vs. $M = 7.78$, $p = .015$). The Cohen’s $d$ test for repeated measures indicated a medium effect size corresponding to .50. The absolute change in the SIRI-R post-training mean score was +1.25 with a relative increase in the post-training mean score of 19.1%. Likewise, the increase in the mean score at the three-month follow-up indicated a statistically significant increase from the pre-training score ($M = 6.53$ vs. $M = 8.32$, $p = .005$) and the medium effect size was maintained (Cohen’s $d = .74$). The absolute change in the SIRI-R follow-up mean score was +1.79 with a relative increase in the follow-up mean score of 27.4%.

For Transportation Supervisors, as well there was a statistically significant increase in the post-training mean score when compared to the pre-training score ($M = 5.57$ vs. $M = 7.26$, $p < .001$) with a medium effect size (Cohen’s $d = .73$). The absolute change in the SIRI-R post-training mean score was +1.69 with a relative increase in the post-training of 30.3%. At the three-month follow-up, the increase in the mean score was statistically significant ($M = 5.57$ vs. $M = 7.78$, $p = .001$) with a large effect size (Cohen’s $d = .93$). The absolute change in the SIRI-R at follow-up was +2.21 with a relative increase in the follow-up score of 39.7%.

To assess the differential immediate and long-term effects of the training on the SIRI-R mean scores of the different occupational groups, a mixed linear fixed-effects analysis for repeated measures was performed using a CS correlation structure. The analysis of variance showed there wasn’t a significant difference between the two occupational groups ($p = .10$). Further examination of post-hoc analyses showed there were no statistically significant differences between the two occupational groups at pre-, post, or at three-month follow-up ($p = .581$, $p = .545$, $p = .587$, respectively).

To assess the impact of occupation on the SIRI-R (outcome measure), secondary repeated measures models were performed adding a term for Time by Occupation (independent variable). Contrast tests permitting post-hoc pair-wise comparisons between time and occupations means were also performed using a Sidak procedure. The analysis of variance
showed that the interaction between the training and occupational group (Training X Occupation) was not significant: $F_{4.206} = 810, p = .520$). Table 26 presents the SIRI-R model based on differences. As demonstrated in Table 26, the greatest difference in post-pre estimated marginal mean scores among the two occupational groups was the difference observed in the pre-post scores of the Transportation Supervisors ($\Delta = 1.75, p = .001$) with the greatest effect ($\Delta/SE$) of 5.65. Similarly, the difference in post-pre estimated marginal mean scores for the Special Constables was statistically significant ($\Delta = 1.39, p = .015, \Delta/SE = 2.84$).

Table 24: Comparison of Proportions of SIRI-R Correct Responses of the Occupational Groups

<table>
<thead>
<tr>
<th>Occupational Group</th>
<th>N</th>
<th>SIRI-R Correct Responses</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$\leq 6/12$</td>
<td>$&gt; 6/12$</td>
<td></td>
</tr>
<tr>
<td>Special Constables</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-training</td>
<td>19</td>
<td>47.40%</td>
<td>52.60%</td>
<td></td>
</tr>
<tr>
<td>Post-training</td>
<td>22</td>
<td>40.90%</td>
<td>59.10%</td>
<td>$^a$</td>
</tr>
<tr>
<td>Follow-up</td>
<td>21</td>
<td>27.30%</td>
<td>72.70%</td>
<td>$^b$</td>
</tr>
<tr>
<td>Transportation Supervisors</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-training</td>
<td>52</td>
<td>61.70%</td>
<td>38.30%</td>
<td></td>
</tr>
<tr>
<td>Post-training</td>
<td>52</td>
<td>30.80%</td>
<td>69.20%</td>
<td>$^c$</td>
</tr>
<tr>
<td>Follow-up</td>
<td>44</td>
<td>35.50%</td>
<td>64.50%</td>
<td>$^d$</td>
</tr>
</tbody>
</table>

$a$ = Comparison pre-post Sidak procedure $p = .591$

$b$ = Comparison pre-follow-up Sidak procedure $p = .141$

$c$ = Comparison pre-post Sidak procedure $p < .001$

$d$ = Comparison pre-follow-up Sidak procedure $p = .009$

Likewise, at the three-month follow-up, Transportation Supervisors continued to lead with the greatest difference in follow-up-pre estimated marginal mean scores ($\Delta = 2.01, p = .001$) and effect ($\Delta/SE = 5.15$), followed by Special Constables ($\Delta = 1.82, p = .003, \Delta/SE = 3.31$).
Table 25: SIRI-R Raw Means by Occupational Group

<table>
<thead>
<tr>
<th>Occupational Group</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>P value&lt;sup&gt;c&lt;/sup&gt;</th>
<th>Effect Size&lt;sup&gt;d&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Special Constables</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-training</td>
<td>20</td>
<td>6.53</td>
<td>2.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-training</td>
<td>23</td>
<td>7.78</td>
<td>2.2</td>
<td>.018&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.50&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Follow-up</td>
<td>21</td>
<td>8.32</td>
<td>2.0</td>
<td>.05&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.74&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Transportation Supervisors</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-training</td>
<td>52</td>
<td>5.57</td>
<td>2.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-training</td>
<td>56</td>
<td>7.26</td>
<td>2.3</td>
<td>.001&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.73&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Follow-up</td>
<td>32</td>
<td>7.78</td>
<td>2.4</td>
<td>.001&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.93&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

<sup>a</sup> = Pre-post comparisons  
<sup>b</sup> = Pre-follow-up comparisons  
<sup>c</sup> = Based on post hoc pairwise comparison using Sidak procedures  
<sup>d</sup> = Based on Cohen’s d for repeated measures

Several additional mixed-model analyses were conducted to ascertain whether an interaction effect existed between the training and the following variables: age, gender, years of employment at the TTC, and experience with suicidal behaviour. None of the aforementioned analyses yielded statistically significant results. Separate mixed-Model analyses were performed with linked and unlinked scores of the SPQ, SOQ-R, IKT-R and SIRI-R. These analyses were conducted to buttress the decision to analyze the data from the complete sample rather than limit analyses to those participants with linked pre-, post-, and follow-up data.
### Table 26: SIRI-R Model Based on Differences

<table>
<thead>
<tr>
<th>Occupational Group</th>
<th>Difference</th>
<th>SE</th>
<th>Effect (D/SE)</th>
<th>DF</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Special Constables</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-Pre</td>
<td>1.39</td>
<td>0.49</td>
<td>2.84</td>
<td>181</td>
<td>0.015</td>
</tr>
<tr>
<td>FU-Pre</td>
<td>1.82</td>
<td>0.55</td>
<td>3.31</td>
<td>237</td>
<td>0.003</td>
</tr>
<tr>
<td>FU-Post</td>
<td>0.43</td>
<td>0.52</td>
<td>0.83</td>
<td>240</td>
<td>0.79</td>
</tr>
<tr>
<td>Transportation Supervisors</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-Pre</td>
<td>1.75</td>
<td>0.31</td>
<td>5.65</td>
<td>179</td>
<td>0.001</td>
</tr>
<tr>
<td>FU-Pre</td>
<td>2.01</td>
<td>0.39</td>
<td>5.15</td>
<td>220</td>
<td>0.001</td>
</tr>
<tr>
<td>FU-Post</td>
<td>0.26</td>
<td>0.38</td>
<td>0.68</td>
<td>223</td>
<td>0.88</td>
</tr>
</tbody>
</table>

As Table C in Appendix N demonstrates, the findings of the linked scores correspond to those reported for the whole sample; thus, limiting the study analyses to the linked data would have not generated divergent or more exact results than those reported in this chapter.
CHAPTER SIX
QUALITATIVE RESULTS

The results of the second phase of the study are presented in this chapter. In adherence with the analytical framework, the presentation of findings in this chapter follows the TKM and the qualitative findings are presented sequentially with the analyses organized in separate subsections following the analytical framework levels in the following order: Reactions, Learning, and Behaviour.

Reactions to the Training (TKM Level I)

One of the aims of the qualitative component was to identify a number of key themes that were likely to have affected participants’ overall reactions to the safeTALK training.

Central to evaluating participants’ reactions to the safeTALK training is an analysis of the extent of their overall satisfaction with the training they received and of the factors that helped shape their assessment of the training. Several themes emerged from the study participants’ reactions to the safeTALK training: perception of usefulness and relevance of the training, the content and format of the safeTALK workshop, the skill training methods, and perceptions of trainers’ performance. Overall satisfaction with safeTALK can be viewed as a global attitudinal construct that captures individual general attitudes towards learning and evaluation of the training they received. Participants’ perceptions of safeTALK training relevance to their work and their satisfaction with the program quality corresponded to Level I in the analytical framework (reactions).

The initiative taken by the TTC to implement an educational suicide prevention program for employees was praised by all study participants. It was seen as a “positive thing to make employees more aware, it was good, especially dealing with the individuals we do in our capacity as Supervisors,” explained one of the study participants; however, discussing suicide, he acknowledged “might feel uncomfortable” for some employees (Transportation Supervisors #6). “It was a good move to implement the program, it was long overdue,” said one of the Transportation Supervisors,
We are dealing with suicides not on a daily basis but at least directly or indirectly on a monthly basis…look at all the people it actually affects, not only the operator, the people that are directly involved in the actual suicide but also the emergency personnel that have to go there. (Transportation Supervisor #3).

The program was considered as “beneficial” it increased participants’ awareness of suicide intent of patrons and participants’ role in preventing subway suicide. Study participants expressed their support for the gatekeeper program initiative and suggested the program should be extended to all front-line TTC employees, subway operators, bus drivers, supervisors, janitors, collectors…anyone that may be in a position where they are dealing with somebody in the subway. (Special Constable #21).

The overall reactions to the safeTALK workshops were very positive and the majority of study participants (63.3%, 19/30) agreed that taking part in the safeTALK training was worthwhile and extremely beneficial to them and in many instances it provided a new perspective on suicidal behaviour and suicide prevention. “I was satisfied, I left there with a whole different perspective on things” articulated one of the Transportation Supervisors (Transportation Supervisors #5).

Study participants found the training was enjoyable: “it was like an open session” (Transportation Supervisor #4) and it provided a venue where emotions could “vent” and allowed participants to “throw some of the frustrations” (Transportation Supervisor #2) engendered by their experiences with subway suicides, and the effect they have on their lives. One of the Transportation Supervisors expressed his hope that management would decide to maintain and incorporate the suicide prevention workshop in the various mandatory training programs for TTC personnel. Based on his personal experience at the TTC, he was wary. He has seen many programs implemented and terminated or suspended and he expressed his skepticism regarding the endurance of the gatekeeper program:

I hope it doesn’t go away. I hope it just doesn’t, like a few other things that I have seen in my twenty years…programs come and programs fade away…and they are lost in the woodwork…whether it continues to be this program, or it’s a program of similar nature is necessary. (Transportation Supervisor #2).
Reactions: Relevance and Usefulness Aspect of the safeTALK Training

Important factors that influenced participants’ reactions to the training were their perceptions of the usefulness and relevance of the safeTALK training to their work; particularly, the extent to which participants perceived the training had increased their knowledge, and enhanced their skills and facilitated early identification and intervention with suicidal patrons; moreover, the extent study participants perceived they improved their general efficacy in preventing suicide and saving lives.

Transportation Supervisors and Special Constables had contradictory views of the training’s utility and relevance to their work. Overall, Transportation Supervisors found “every aspect of the training relevant” to their work (Transportation Supervisor #2), and, in general, felt the safeTALK workshop increased their awareness and enhanced their knowledge about suicide prevention and the suicidal individuals. In general they felt “there was no information taught that was useless.” (Transportation Supervisor #5). They acknowledged the training enhanced their ability to identify distressed patrons and led to an overall increase in vigilance, to being more “alert to people that looked like they may be contemplating suicide” (Transportation Supervisor #6) and they “gained a little more insight into how to deal with these patrons.” The safeTALK workshop elucidated and highlighted behaviours that may be indicative of distress and of potential suicide risk, warning signs they were unaware existed and were never trained to recognize or educated in how to intervene. As one of the Transportation Supervisors acknowledged,

It was relevant to me; it pointed out stuff that I wasn’t aware of uh and things to look for to identifying people at risk...I could somewhat relate to and I’ve used in my day to day work. (Transportation Supervisor #13).

The suicide warning signs indicate heightened risk for suicide and set a temporal relationship between pre-suicide attempt behaviours and actual suicide attempts. Sensitizing participants to warning signs as valid indicators of suicide danger, as one Transportation Supervisor succinctly observed: “was most relevant to my work...to look for signs of people that may want to commit suicide ’cause normally you’re not attuned to that when you’re in our environment.” (Transportation Supervisor #15).

There were mixed feelings expressed by Special Constables about the training’s relevance to their job. While they felt the training was relevant and beneficial for Transportation
Supervisors, they perceived the workshop provided very elementary training, did not apply to them and the content did not reflect the legal authority Special Constables have on TTC property. Additionally, they felt the content of the workshop did not consider their extensive training, their rich and diverse experiences, and their abilities to intervene in crisis situations. As one of the Special Constables noted: “the training was good. It was very extensive, very extensive, there was a lot of training”; however, he had reservations about the content of the workshop, he found there were:

A lot of things lacking specifically for the law enforcement aspect, the job we do. The training specifically wasn’t really relevant to my job I uh it was more relevant I think to supervisory personnel…It didn’t really cover our responsibilities as peace officers having the power to make an apprehension…um, but in saying that I’m a bit uh maybe I’m a bit biased. (Special Constable #21).

His response and sentiments were echoed by several other Special Constables who also found safeTALK did not meet their training needs. In addition to being very basic training, it did not take into account their previous mental health training, their familiarity with the Mental Health Act, and their extensive experience enforcing the Mental Health Act:

I found that the uh training that was available there was a somewhat of a very basic level, taking a person from the ground, right, from the bottom floor and I found that my particular experience and exposure and the type of uh situations that I’ve uh coped with was somewhat more sophisticated if you will or more advanced…(Special Constable #1).

One Special Constable summed up the experience of attending the workshop:

I think the training was relevant to the civilian staff here at the TTC. It was a great introduction, it was well taught, all that stuff, but it was definitely, not for us, I mean it was, I don’t want to say, it but it was a waste of time. I truly think it was. (Special Constable #29).

Reactions to Workshop Format

Overall, the safeTALK workshop was well received by the Transportation Supervisors and very little criticism of the content, format or delivery was expressed by them: “It was all very well laid out and very well uh very well handled.” (Transportation Supervisor #6). The teaching methods that promoted interaction among workshop participants were liked by the Transportation Supervisors who found the training format to be “Very good. No ‘talking heads’, interactive, very open…” (Transportation Supervisor #7).
No I think it was good because there’s, no matter what job you’re doing everyday is finding out something new about something or somebody else job. There are a lot of people out there that still think that we’re security they don’t know that we can do the full spectrum of apprehension and taking them to the hospital. So it helps, I think it helps a lot… it’s good in more ways because as long as there’s a good working relationship with everybody. (Special Constable #30).

Nonetheless, there were Special Constables who voiced criticism of the training format and delivery. Those Special Constables believed the current format of safeTALK to be inadequate and did not meet their training requirements. Specifically, their issues were two-fold: 1. the same workshop being offered to Transportation Supervisors and Special Constables; and 2. having a mix of occupational groups attending the same workshop. While they felt “the program was very helpful and very informative for Supervisors” having Special Constables and Transportation Supervisors in the same training session “was like mixing first graders and eighth-graders in the same class” (Special Constable #1). It was opined that Special Constables needed more specialized training that included a review of specific disorders associated with suicide and law enforcement aspects of the Mental Health Act, whereas the emphasis of the suicide prevention training for Transportation Supervisors should be on skill training enhancing their ability to communicate and intervene. The rationale for having separate training was provided by one of the Special Constables “Because our jobs are so completely different” (Special Constable #1).

For some study participants learning was enriched by the interdisciplinary approach. Sharing experiences with participants from different occupational backgrounds helped establish, for some study participants, a wider network of support within the TTC, and afforded a deeper understanding and appreciation for the different role each occupational group has when confronted with the suicide of a patron or a suicide incident. As one Transportation Supervisor reflected:

I thought that that was helpful because I was able to hear their views, I was able to listen what they had to say, to listen to their opinions… ’cause suicide is you know is kept so secret. (Transportation Supervisor # 2).
This viewpoint was not shared by some of the Special Constables who felt that having a combined training for both occupational groups i.e. Special Constables and Transportation Supervisors made the training irrelevant and unhelpful.

There were also mixed opinions about the length of the workshop. The workshop was “a bit too short” and could have benefited from a longer discussion on the resources that are available for suicidal patrons, remarked one of the study participants (Special Constable #26). Conversely, others found the training “too long and too dragged out” and felt it did not require “a full day” (Transportation Supervisor #13) and could have been shortened and delivered in “about four hours of good concentrated training time.” (Transportation Supervisor #5). One of the participants commented:

I don’t think the program needs to be a full day course…I am fearful that a lot of people are just tuning out and they are just sitting there for the full day. So if it was a half-day thing, if you could integrate into a half day, I think it would be more potent. (Transportation Supervisor #15).

While he understood that “presentations that get the group involved and participating” and interactive methods of teaching, such as role playing required time to organize and execute, he was “fearful that not a lot of people took it too seriously because of the duration...” (Transportation Supervisor #15).

Reactions to Workshop’s Content

The vast majority of study participants felt the behavioural warning signs were a valuable component of the workshop and found the list of behavioural cues was “comprehensive” (Transportation Supervisor #13) and “alerted” them to “what people that may be contemplating suicide look like” (Transportation Supervisor #8). It provided most of them with:

An insight into how people are withdrawn and uh and how to recognize these people…to identify that perhaps they’re a little distraught, uh maybe uh withdrawn for a reason. (Transportation Supervisor #6).

For those who were familiar with some of the behavioural cues, the training expanded their knowledge and brought it to a “higher level” (Transportation Supervisor #3), as one of the Transportation Supervisors explained: “I’ve always been good with reading body language, but this course did help me identify people that may be contemplating or may be distraught”
(Transportation Supervisor #5). Regardless of the number of years of experience, the behavioural cues, “how to” detect patrons who might be distraught and at risk of suicide were considered as a “good” (Special Constable #30).

Other criticism of content was made by Transportation Supervisors who found the information component at the start of the workshop irrelevant to the training and have indicated that knowing that “one in twenty thinks about suicide that wasn’t completely helpful or relevant” (Transportation Supervisor #11). A second Transportation Supervisor felt “the emphasis that everybody is at risk, that everybody knows somebody or is associated with somebody who’s thought about suicide” was seen as a point of view that did not require discussion, nor was it considered as helpful in raising awareness about suicide:

Suicide is a solution for them and I always believed that, I’ve always had an understanding of that, sure I don’t understand why they choose to jump in front of a train but I do understand and the fact that you know some people make that choice. (Transportation Supervisor #13).

Reactions to Experiential Learning: Role Playing

The mastery of skills relies on the opportunity to practice new skills and have those reinforced over time. Skills acquisition is an integral component in any program targeting improvement in the communication process. Much evidence suggests that the most important component of communication skills training is experiential learning through role play (Kurtz, Silverman & Draper, 1998). Role playing is a method of portraying human interaction in imaginary situations in such a manner that realistic behavior is elicited. Role playing provides trainees with opportunities to observe, experience, and practice actual behavior in contexts somewhat similar to reality. The rationale for role playing starts from the conviction that the aim of training is not solely to transmit facts or viewpoints, but to help the trainee translate knowledge so that it becomes meaningful in his/her own experience. The objective of role playing is to increase the trainee’s awareness of the implication of his/her actions, and the actions of other trainees toward him/her. The purpose is also for the trainee to become skillful in assessing and taking actions in ongoing situations. One requirement for the development of this awareness is opportunity for the trainee to actually experience functioning in realistic situations. Role playing provides this opportunity and consequently increases confidence.
There were mixed reactions to the role-play component of the training. Some study participants found the incorporation of interactive methods of teaching and the opportunity to practice and rehearse their communication skills through role playing beneficial. As one Transportation Supervisor stated: “the role playing was quite positive for me; it was a very important component.” (Transportation Supervisor #6). Those study participants who had mentioned role playing as a positive and useful learning experience had stated the role playing made them feel more confident and provided them with hands-on-experience. “It was practical yeah, because we don’t often get a chance to practice that on the job.” (Transportation Supervisor #10). Playing the role of someone in need of assistance gave rise to various emotions the participants believed were experienced by someone considering suicide. The role playing “put you in their shoes” (Transportation Supervisor #10) and “it gave me a chance to see how other people react to what I was saying and gained more insight into how to deal with traumatized individuals,” explained Transportation Supervisor (# 6). Several other study participants found the role-playing emotionally difficult. Empathizing with the suicidal patron and placing themselves “into that position” proved to be difficult for several Transportation Supervisors, especially when they had to confront their own emotions and personal beliefs about suicide. They found the experience “upsetting”… “Because sometimes I can’t figure out what would make somebody like...what could be so bad to make you think like that...” (Transportation Supervisor #4). Suicide drew forth an array of intermingled emotional responses. One of the Transportation Supervisors confessed that while he found role playing allowed him to gain insight and to be reflective about his own attitudes toward suicide, he was conflicted and “found it quite hard to empathize with a suicidal patron.” (Transportation Supervisor #10).

Role playing as a teaching method was also criticized by a small group of the Special Constables who felt that although role playing “was good” and beneficial, it was considered to be more appropriate for junior employees who lacked the experience of dealing with suicidal patrons. One of the Special Constables felt the role playing component did not accurately depict incidents which Special Constables encountered from time to time. One such incident involved intervention with “the regulars,” patrons who are mentally ill and who are well known to Special Constables from multiple past encounters. Intervention in those instances did not involve assessment of suicide risk as they were deemed as attention seeking behaviour. It was suggested that role playing was more valuable to the trainers than the workshop participants as it exposed
the trainers to incidents they did not have knowledge of: “I think it open your eyes...It may have taught you people more than it did us because we see this on a regular basis and we deal with these people.” (Special Constable #26).

Reactions to Video Presentation

The videos presented in the safeTALK workshops were developed after extensive consultation with TTC personnel and clinicians to best exemplify behaviour that may be consistent with distress and suicide risk. The reactions to the video presentations were positive. Participants remarked they found the videos helpful and beneficial. It provided them with visual examples of a suicidal patrons’ demeanour and gave tangible examples of behavioural and verbal cues. As one Transportation Supervisor stated: “the videos showed you how they might appear” (Transportation Supervisor # 10) and thus the videos alerted participants to the subtleties of the warning signs and taught participants how not to miss or dismiss potential cues or invitations to help.

Reactions to Presentation Techniques: Group Discussions

Group discussions were effective at different points in time in the process of training. They were effective at the start of the training session to alleviate polarization of opinions and early perceptual problems. They were also effective in fostering an environment conducive to learning, and they enhanced learning in both the affective and cognitive domains. The objectives of group discussions were two-fold: (a) for the trainees to acquire new insights by hearing different perspectives and by having their ideas critiqued, and (b) trainees learn and commit to new behaviors from group discussions. Group discussions with ground rules were especially valuable and effective when the content of the training touched on controversial issues and allowed for candid discourse on a subject many at the TTC believed to be an unmentionable topic.

The group discussions provided a non-threatening venue for all participants to express opinions on deeply-held personal beliefs about suicide and encouraged mutual respect and constructive listening. They also encouraged the participants to exchange their own experiences, thereby making learning more active and less isolating.
The reactions to the group discussions were mixed. Some study participants found the group discussions were enlightening and being able to hear the other participants’ perspectives and opinions were insightful. These study participants viewed the experience positively and felt they benefited from having a mixed class which was comprised of trainees with diverse occupational backgrounds, i.e.: Assistant Superintendents, Superintendents, Chief Supervisors, Route Supervisors, and Special Constables. In their opinion it allowed for candid discourse on a subject many at the TTC believed to be taboo.

As one special Constable clarified:

It wasn’t so much the training, but… the group discussions, I remember different points of view from different people whether it was a Special Constable or uh TTC Supervisor and it was very interesting to hear their perspective on a lot of this stuff. Everyone brought their own experience to the table and we all had a share and whether we took something away from it we probably made us think how we react to calls. (Special Constable # 19).

In general, training sessions started with an introduction discussion. The introductory discussion of the workshop was designed to establish a common ground between the trainer and trainee, to capture and hold attention, to outline the workshop and relate it to the overall training, to point out benefits to the trainees, and to lead the trainees into the body of the workshop. It usually contains a segment that provides a general statement of the workshop content, target population, the rationale for the training, and appropriate motivation to gain the trainee's attention. The safeTALK workshop commenced with a general discussion about suicide myths and facts, beliefs and attitudes.

Some Transportation Supervisors cited the suicide facts as irrelevant to their work. Knowledge regarding the prevalence of suicidal thoughts in the general population was not deemed, by some, as helpful or as a factor that can contribute to their sensitization. As one Transportation Supervisor shared: “I don’t need that sensitivity when it comes to the subject… I don’t need to be told that everyone thinks about it and stuff like that…” (Transportation Supervisor #11). Another Transportation Supervisor thought the presentation of those facts was to a certain extent manipulative:

They just say everyone thinks about it, you know if it’s not you, well then it might be your brother, it might be your cousin or something, so they try to hit home with it. I
think they try to win people on-side with the idea of suicide…(Transportation Supervisor #13).

**Reactions to Trainers’ Mastery of Knowledge**

Participants’ overall satisfaction with the training is positively associated with their appraisal of the trainers’ performance. The trainers’ mastery of topics and delivery performance also moderates the relationship between the perceived relevance and usefulness of the training and training satisfaction (Giangreco, Sebastiano & Peccei, 2009). Thus the perceived relevance of the training has a positive effect on training satisfaction when the study participants perceive the trainers’ performance to be high.

The issue of the trainers’ level of preparedness and familiarity with the responsibilities, duties and legal powers of Special Constables was broached by some of the study participants. Special Constables alluded to the fact that the trainers were not aware of the legal powers Special Constables have and their ability to apprehend patrons who are at risk of suicide and to transport them to a hospital emergency department: “when outside sources come into the TTC, they don’t know enough about the department” (Special Constable #30). The trainers were not “really up to speed with what Special Constables dealt with, our wealth of experience of everyone in the room…so they were a little caught off guard.” (Special Constable #19) and subsequently increased the trainers awareness of the extent of Special Constables powers, experience and skills and contributed to a new appreciation of their abilities. “It may have taught you people more than it did us because we see this on a regular basis and we deal with these people,” remarked one of the Special Constables (Special Constable # 26).

**Learning (TKM Level II)**

The immediate aim of the safeTALK training was to enhance suicide intervention knowledge and skills of trainees, with the intention that they can recognize and respond to a patron at risk. This section of the chapter examines whether and how safeTALK enhanced study participants’ knowledge and skills, and whether the training changed their attitudes toward suicide and suicide prevention.
Warning Signs

Implicit in the general suicide education approach is recognition of the difficulty of determining who, among millions of healthy TTC patrons, is truly at high risk of suicide. Warning signs are an early indication that a patron might be at a heightened risk of imminent suicide. Warning signs may be a cry for help and they can provide an opportunity to intervene and potentially prevent the suicide from happening. To facilitate the identification of patrons at risk for suicide, a list of 20 warning signs was developed after extensive consultation with key informants who provided their expert opinions on warning signs exhibited by patrons who have attempted or suicide in the Toronto subway system. The warning signs were incorporated by \textit{LivingWorks} into vignettes developed specifically for the TTC and integrated in the \textit{safeTALK} workshop.

During the development process every attempt was made for the list to: a) accurately represent behaviours associated with distress and suicidality; b) be comprehensive; c) unambiguously portray the behavioural cues; and d) be beneficial for the identification of distressed patrons who might be at risk for suicide. The comprehensiveness and clarity of the list was validated with stakeholder groups: mental health personnel working with the St. Michael’s Hospital Psychiatric Emergency Service, TTC Special Constables, Mobile and Surface Supervisors, and Train Operators.

Several study participants reported gaining knowledge of the warning signs. Numerous Transportation Supervisors considered the list of behavioural warning signs new and novel information and as one Transportation Supervisor stated: “\textit{I never learned that before}” (Transportation Supervisor #11). The information was valuable and beneficial and facilitated the identification of spotting potential suicidal patrons when we are in the subway to look for more signs of people that may want to commit suicide, ‘cause normally you are not attuned to that when you are in our environment, in the subway…it was an eye opener to sort of look for certain signs…it sort of open your perceptions a bit to make you more aware…(Transportation Supervisor #8).

\textit{“I was always good with reading body language, but this course did help me identify people who that may be contemplating suicide,“} noted one of the Transportation Supervisors (Transportation Supervisor #3). While another Transportation Supervisor stated: “\textit{it alerted me to what people that may be contemplating suicide would look like}” (Transportation Supervisor #10).
I thought that was quite good. They gave you an insight into how people are withdrawn and how to recognize these people, whereas before you might think they were just standing there for no reason...but now they trained you to identify that perhaps they are distraught...and uh maybe withdrawn for some reason...and how their overall presentation is affected by their mood. (Transportation Supervisor #21).

Not all Special Constables found the warning signs as important or beneficial to identifying a potentially suicidal patron. One Special Constable explained: “We are basically responding to calls. We are not really out there looking because we don’t have the manpower” (Special Constable #21).

A few study participants felt the workshop did not offer them any new knowledge. One of the Transportation Supervisors observed: “they didn’t offer you anything new, any new insights. I was aware of the warning signs.” (Transportation Supervisor #5). It was a view that was shared by one of the Special Constables who claimed:

I’m pretty much aware of the warning signs, I really am, uh I see them all the time and I’m aware of the warning signs, I know what they are...there are a lot of clues...but uh um I can tell. I can tell a suicidal person. (Special Constable #1).

**The Effects of the Training on Skills Acquisition**

The overall concept of suicide prevention is to intervene at some point in the process of an individual’s path towards killing him/herself. Intervention skills consist of concepts of empathy, active listening, communication skills, a nonjudgmental attitude, and appropriate verbal interventions.

Skill refers to the learned capacity to perform specific behaviours, to complete tasks or activities. Skills training enhances the capacity for critical thinking, and teaches participants to apply these cognitive processes to guide their behaviours. Skills training also encourage the development of personal insight and the ability to recognize when emotional reactions are negatively affecting objectivity and the accuracy of judgment and conclusions.

Many of the interviewed participants cited as valuable the opportunity to learn new techniques on how to approach suicidal patrons and how to engage them in a direct and open discussion regarding their suicide intent. As one Transportation Supervisor admitted, “I think for
the most part a lot of us wouldn’t know that if we were not told.” (Transportation Supervisor #15).

**Engaging in a Direct and Open Communication about Suicide**

In any communication between two people there is a margin of error between what the speaker intends and what the listener hears and understands. Study participants were taught that suicide warning signs are at once observable behaviours, pre-suicide indicators and danger alarms that provide the unique, and sometimes the only, opportunity to intervene in a developing suicide crisis. Suspected warning signs can be validated or invalidated by asking one or more clarifying questions, and there are no negative consequences in learning that someone is not suicidal. Asking clarifying questions helps decode indirect but detectable suicide communication signals. Study participants were taught that at some point in the interaction with a distressed patron a straightforward question regarding suicidal intent is required, e.g.: “Are you thinking of killing yourself?” The clarifying question creates an active dialogue between gatekeepers and patron by which the true meaning of the statement can be fully, correctly, and completely understood. Learning to ask the clarifying question is a key to saving a life. A confirmation of meaning can then be followed by a helpful intervention, referral, and treatment.

The type of skills study participants felt they acquired through the *safeTALK* training was their communication skills. Study participant felt the training provided them with “guidelines” (Special Constable #25) and with “the tool to introduce the dialogue and to engage with the person” (Transportation Supervisor #13). Additionally, they also learned “not to skirt the issue,” and not to try to “beat around the bush” (Special Constable #29), but rather “just go straight to the point.” (Special Constable #26). “They sort of gave us the start to questioning people, to probe them regarding their intentions,” explained a Transportation Supervisor (Transportation Supervisor #4). Being taught to engage in direct suicide communication and learning the meaning of oblique suicidal communication —whether verbal or behavioural— provides a conversational context in which an active dialogue between speaker and listener can be created:

The approachability. A better way to approach the individuals and to be able to talk about it, very openly, without putting myself in any jeopardy or danger. The program taught me um to be very direct and when we approach people and I never did that before…go right to the person and say “are you thinking of committing suicide? (Transportation Supervisor #3).
Prior to attending the training, observing behaviours that were indicative of distress and potential suicide risk produced in many of the study participants, apprehension, fear and misplaced sense of responsibility for the potential possible actions of suicidal patrons. And while participants were “compassionate” and “gentle” with distraught patrons (Special Constable #21), they were reluctant and afraid to ask distressed patrons whether they were suicidal. These strong emotional reactions to a possible suicide crisis inhibited a helpful response. For some study participants the reluctance to approach a patron stemmed from their fear of negative consequences if the suspected warning signs were not validated by the patron. A second concern which hampered a helpful response was their fear of placing themselves in a vulnerable and unsafe situation on the platform exposed to potential “jeopardy or danger” (Transportation Supervisor #3). Additionally, study participants had reservations about approaching and questioning a patron who appeared distraught as they feared such action may lead to an angry reaction from the patron followed by a “call to the police and filling a complaint about being harassed.” (Transportation Supervisor #15). Consequently, as another Transportation Supervisor explained: “I wouldn’t approach suicidal patron…I would call it in and watch or follow them, but I would not approach them.” (Transportation Supervisor #6).

A number of participants spoke about their reluctance to ask a distressed patron if they were thinking about suicide. Their hesitation was engendered by fear the question may trigger thoughts of suicide or suicidal behaviour or heighten existing suicidal feeling. They feared engaging in direct dialogue about suicidality might be “prompting them to react” (Transportation Supervisor #4) or “forcing them to do it.” (Transportation Supervisor #3). The training helped demystify myths and misconceptions they had about suicide and the suicidal patron:

It was an eye opener. I always had fear of planting the idea, sending the person over the edge…and I’ll have to take responsibility for this person jumping in front of the train…(Transportation Supervisor #2).

I was always afraid that I was going to trigger them to say “Oh my God there are people here to prevent me from doing something, now I am forced to do it right now”. (Transportation Supervisor #3).

Lack of training and being uncertain how to respond in the event the patron endorsed suicidal thoughts made most participants apprehensive, and thus, in most instances, they chose not to ask about suicidal ideation. The training, as one participant stated, provided knowledge,
communication and intervention skills which, he felt, can be put into practical use in an intervention situation:

I think that the whole combination of things...you can talk instead of shying away from them or just calling for help. You can actually approach them, maybe instead of just grabbing them and holding them, you can relate to them, maybe make them feel less vulnerable... (Transportation Supervisor #10).

Although conceptually, for Special Constables, the training did not present new knowledge “it helped reinforce” (Special Constable #20) their existing knowledge acquired from earlier training, in their years of experience working at the TTC, and previous experience in law enforcement-related professions. As one Special Constable explained:

I think in my case, I am thinking about me with 22 years of experience plus having worked in the military...the how to; I mean it’s something you learn over time as well. How to look at an individual and by looking at their body language and the way they carry themselves, make a decision whether they are at risk...as well the way to talk to them and to communicate and ask. (Special Constable #24).

In contrast to study participants who found the communication and intervention skills valuable, some Special Constables felt

The intervention techniques that were taught were more for that of a civilian dealing with someone who is in crisis versus a peace officer or police officer...I don’t want to disrespect the training because like I said it was excellent for the civilian staff. (Special Constable #29).

For some Special Constables being taught that it was necessary to take time to develop rapport with the distressed patron was unrealistic. When they are called to intervene with a potentially suicidal patron their main concern is to secure the platform, to ensure the safety of other patrons, and not to place themselves at risk. As one Special Constable explained: “On the train platform with the train coming I don’t have the time to be nice. We don’t have the luxury of being nice. You tell me you are going to commit suicide, you are under arrest, period!” (Special Constable #27).

We are not counsellors so and I find with that aspect was a little bit too in depth for us. We need more a first response, quick action type involvement and as result maybe the training should have been more gear to that. (Special Constable #25).

It was a view that was shared by one of the Transportation Supervisors (#3) who explained:
The length of time I actually have to assess the situation. I only have a very short time frame to assess the situation and sometimes you need a lot more time to assess. One of the first things I need to do is secure safety for myself, the platform, other patrons, so that requires quite a bit of uh actions on my part. (Transportation Supervisors #3).

Transportation Supervisors dispatched to a platform are on their own and one of the first priorities is the need to secure safety on the platform for themselves and for patrons. Conversely, Special Constables are always dispatched in pairs to investigate a potential suicide attempt.

Attitudes

Attitudes are favourable or unfavourable evaluative reactions towards something or someone, exhibited in one’s beliefs, feelings, or intended behaviours (Myers, 1990). Attitudes have three dimensions: a) affective component which consists of the emotional response that expresses an individual’s degree of preference to the objects of the attitude; 2) behavioural component which consists of action(s) or observable behavioral tendencies that are the result of an attitude object; and 3) cognitive component which constitutes the individual's thoughts and beliefs about the attitude object. Positive attitudes toward suicide prevention may influence the willingness to intervene in a crisis situation and may also serve as strengthening and enabling factors in suicide prevention. Attitudes are expected to change as a function of experience and through persuasion (Tierney, 1994; 1988; MacDonald, 1999). One of the key modules in the safeTALK workshop was a group discussion of attitudes to suicide. The purpose of the discussion is to help participants explore their own attitudes and those of other participants toward suicide and suicide prevention and to reflect on how their attitudes might affect their willingness to intervene, and the intervention process (Tierney, 1988).

Most study participants (83%, 25/30) indicated the training did not influence their attitudes toward suicide; nor did they feel their attitudes toward suicide were altered as a result of the training. Study participants attributed their attitudes toward suicide to

Personal beliefs…my attitude towards suicide are strongly instilled in me…dealing with suicides and the clean ups and the aftermath…and the actual suffering of the individuals that are involved…the suicide prevention program is not sufficient to change my attitude…it help me better understand the suicidal patron. (Transportation Supervisor #3).
Although some study participants confided that their attitudes have not changed, the training; however, facilitated a better understanding of “the possible frame of mind of an individual…what actually causes them to want to end it all.” (Transportation Supervisor #4) and increased their “empathy level,”

It made me understand why they may be contemplating suicide…prior to the course, inside of me I would be angry that they would want to do this…to disrupt so many people’s lives, to interrupt my job, and make my job more difficult…how could they! But now I tend to be more sympathetic, more concern for their well being. (Transportation Supervisor #5).

Others felt their personal attitudes toward suicide and views about suicide prevention were not challenged by the training as they did not differ from those promoted in the training, “they were along the same lines.” (Transportation Supervisor #8). As one Special Constable explained, “they were trying to make us be more sympathetic to these people and trying to show that they are in need and I generally carry that feeling…it was already part of me.” (Special Constable #24).

Others attributed the lack of change in their attitudes toward suicide to their personal experience with mental illness and suicide. Several study participants divulged they had close family members who were diagnosed with bipolar disorder, depression and who have attempted suicide; consequently they had first-hand experience with “depression and the desperation,” they have been “exposed to suicide” and understood “the psychology of depression” (Transportation Supervisor #15) and “the personal tragedy behind suicide” (Special Constable #21). “We all walk in different shoes” (Special Constable #19) everyone is vulnerable, each person has their own personal tragedies, their own sorrows and anguish “so it can happen to anybody” (Special Constable #19).

Other study participants admitted “I wouldn’t have wanted my attitudes to change” (Transportation Supervisor #10), while another Transportation Supervisor declared

I don’t have a view on suicide. It’s not something that I think about. I feel sorry for the person…there got to be another way besides suicide…there has to be a better way of dealing with whatever problem that person had...(Transportation Supervisor #12).
All affectively based attitudes have three key things in common. They don't result from rational evaluations, they are not governed by logic, and they are often linked to people’s values. Often, affectively based attitudes come from religious and moral beliefs. One Special Constable’s attitudes toward suicide were clearly influenced by his strong religious convictions. He considered suicide “a selfish act...the ultimate sin” and the decedents who have chosen to take their life “in effect have chosen to play God...and they won’t be forgiven for what they have done. God has given life for a reason...life is sacred.” (Special Constable #17).

One of the Special Constables admitted his attitude changed after the training and the realization there were resources available in the community to provide help to suicidal individuals. He felt relief, the burden and responsibility was not his and his colleagues’ to bear alone:

My attitude changed because I realized that there was a lot more help out there. More help than we were aware. We only see one aspect...we take them to a hospital...I felt that there’s some help out there. We’re not alone. We do our end of it, but then somebody else has a responsibility to do a little bit more. (Special Constable #26).

**Behaviour (TKM Level III)**

Kirkpatrick’s third level of outcome, behaviour, refers to the transfer of knowledge, attitudes, and skills acquired in training to behaviour into the work setting. (Kirkpatrick,1960a). The study design did not include methods of obtaining direct measures of the transfer of acquired knowledge and skills to the workplace setting. The evidence, therefore for Kirkpatrick’s third level of outcome is derived from study participants’ reported accounts of the changes they observed in their own alertness, their communication and interactions with distressed patrons, and their perceived competence, confidence, and comfort when intervening.

**Alertness**

Interviewed participants’ perceptions on the effects of the training on their alertness level were mixed. Some study participants felt their alertness level increased since the training and found that they were “paying attention a bit more...just kind of take a second look and wonder what they might be up to” (Transportation Supervisor #13), and reported the training “heightened” their “awareness of what to look for” (Transportation Supervisor #10), what to
watch for and being alert “for unusual behaviour” (Transportation Supervisor #3), “watching for more signs.” (Transportation Supervisor #15). The vigilance and alertness to unusual behaviour continued when they were off duty using the subway system:

When I am off duty at the subway station, um I always watch anyway...I might pick up on something that I hadn’t picked up on before just because of the awareness aspect. (Transportation Supervisor #6.)

One Transportation Supervisor described his behaviour on the platform since he attended the workshop:

I tend to be more alert on the platform...since the course, it was always kind of there...but now I walk the platform and if I see anything out of the ordinary...Well just this morning I had a lady standing here for fifteen minutes and I approached her and asked her if she was okay, if there is anything that I can help her with...I may not have done that prior to the suicide training...(Transportation Supervisor # 3).

Some study participants felt the training helped them develop “a personal and professional deep commitment to trying to save these peoples lives” (Special Constable #22), which subsequently increased their alertness level on the platform. One Transportation Supervisor stated he was “a bit more paranoid of what’s going on...you go ‘oh maybe this guy’ or ‘this person.’” (Transportation Supervisor #4). Another Transportation Supervisor admitted that after the training he was definitely “more alert on the platform” and although he did not believe the training increased his alertness level by “ninety percent,” it did most likely increase it by “maybe twenty, twenty-five percent.” (Transportation Supervisor #4).

Conversely, another Transportation Supervisor felt his level of alertness increased by

One hundred percent. Having gone to the course, yes now I look for it, before I didn’t. I felt well if it’s going to happen it’s going to happen. Now I’m on I’m on my own alert and I tend to be more alert on festive seasons (Transportation Supervisor #12).

Others described the change in their alertness

…paying attention a bit more...You used to go up there and you’d just do your job and you’ll walk on by, people come and people go, and you’re sort of used to it…Now you’re watching and questioning…(Transportation Supervisor #10).
I’m on high alert, and I’m watching, and listening, and analyzing everything because uh this is important business when you’re talking about people’s lives. You can’t make mistakes. (Special Constable #1).

The training increased study participants’ knowledge about the warning signs, provided them with factual information about suicide and suicide prevention, and taught them strategies to communicate with a distressed patron and skills to intervene if the situation warranted intervention. One Transportation Supervisor stated

It has really changed my intervention style. I’ve used it many times since the training. Before the training, I didn’t want to be in someone’s face…I didn’t think of it being effective to walk up to somebody and say ‘do you think you’re suicidal?’ (Transportation Supervisor #13).

In general, Special Constables felt the training did not contribute to an obvious change in their level of alertness nor has their behaviour changed since the training: “I’ve always been on the lookout,” said one of the Special Constables (Special Constable #30), “I’ve always been a bit of a people watcher” (Special Constable #21).

Being highly alert and observant was perceived by some Special Constables to be an indispensable part of their work they do, their training as Special Constables, and their years of experience.

My level of alertness is high, but I also think that’s my job, that’s what I do for a living. I need to be able to identify people who are at risk more quickly than the average citizen. (Special Constable #24).

I am always alert, always checking for unusual behaviour. It was brought by the training I have as a Constable and my years of experience. Whenever I am at the platform I look for people behaving oddly. (Special Constable #1).

**Ability to Identify Distressed Patrons at Risk for Suicide**

Study participants indicated the training had a positive effect on their capacity to recognize suicide risk warning signs and their ability to identify distressed at-risk patrons. Study participants indicated the *safeTALK* workshops heightened their awareness of the warning signs and circumstances and situations where a person might be contemplating suicide. This perception was shared by most interviewed participants. The vast majority (86.2%, 25/29) of the interviewed participants felt the training provided them with the specific knowledge they required to be able to identify distressed patrons who were at risk of suicide. The mean rating for
this item was 4.24 (SD = .78, Mdn. = 4) indicating that the average self rating of ability to identify patrons at-risk of suicide may be described between “able” and “very able”. Special Constables self-rated their ability to identify distressed patrons higher than Transportation Supervisors (M = 4.6 SD = .5 vs. M = 3.8, SD = .9, p = .01). As one Supervisor stated: “I feel that I can identify a lot more someone who wants to hurts themselves or someone that is incapable of making a safe decision for themselves.” (Transportation Supervisor #11).

Competencies

Competency is defined as “a combination of attributes enabling performance of a range of professional tasks to the appropriate standards.” (Gonczi, Hager & Oliver, 1990: 62). Competency is more than just knowledge and skills. It involves the ability to meet complex demands by drawing and mobilizing psychosocial resources (including skills and attitudes) in a particular context. Competencies result from a variety of factors including the individual’s motivation, self-concept, knowledge, and/or skills. (Gonczi et al., 1990:62). Behavioural competencies are the abilities, attitudes, and values required to perform effectively in a job function or role. Behavioural competencies are typically learned and developed through life experience and they refer to the ”How” of performing a job, and they complement technical competencies.

Overall, interviewed participants felt that their knowledge and competencies for suicide interventions were enhanced by the safeTALK experience. They felt they were better able to make direct inquiries regarding the presence of suicidal thoughts and intent and felt better prepared to assess factors related to suicide risk.

Interviewed participants were asked to self-rate their competencies to deal with a distressed, at-risk of suicide patron. The competencies included ability to identify a distressed, at-risk of suicide patron; ability to engage a distress at-risk of suicide patron in direct and open talk about their intent; their sense of competence, comfort, and confidence. These competencies were self-rated separately on a five point scale, which ranged from “not at all” to “much” (able, comfortable, competent, confident) and corresponded to the numbers one to five, respectively.
Perceived Competence to Intervene

There was an overwhelming agreement among interviewed participants that safeTALK training helped them develop more mastery and increase their ability to intervene with suicidal patrons. The vast majority (97%, 28/29) of interviewed participants perceived themselves to have the competence required to intervene with a suicidal patron, and among those, 51.7% (15/29) perceived themselves to be very competent. A mean self-rating of 4.47 was observed for this item, indicating that the interviewed participants’ perception of their competence to intervene with a suicidal patron may be described between “competent” and “very competent.” Special Constables self-rated their competence to be greater than that of Transportation Supervisors (M = 4.6, SD = .5 vs. M = 4.2, SD = .6, p = .042).

Both qualitative and quantitative data shows that after the training participants felt more confident to intervene with a potentially suicidal patron. The increase of confidence was attributed to a greater understanding of the suicidal individual and to the knowledge that the topic of suicide can be broached openly and directly with distressed patrons.

The training increased my confidence in asking directly if they are suicidal. Prior to the training, we used to skirt around the issue without asking them (directly)…We were told not to ask if they are suicidal.” (Transportation Supervisor #10).

A few of the Special Constables reported the training altered their confidence levels; however they were reserved in attributing their overall confidence to the safeTALK training. “I would say somewhat. It just gave me some added information that I could use in dealing with the situation um which I find maybe very practical and successful.” (Special Constable #25). “It was like putting an extra layer of confidence,” declared one of the Special Constables. (Special Constable #29).

Ability to Engage in Direct Communication

The training appears to have impacted safeTALK participants’ ability to engage distressed patrons in direct and open discussion about suicide. Study participants stated the training helped them feel more confident and self-assured and ready to confront a person contemplating suicide. Most interviewed participants (97%, 28/29) felt that as a result of the training they were able to engage patrons at-risk of suicide in direct and open talk about suicide. The mean self-rating observed on this item was 4.4 (SD = .57, Mdn. = 4.5) indicating that the
average self-rating ability to engage in direct and open talk about suicide may be described between “able” and “very able.” There was no statistically significant difference between the responses of Special Constables and Transportation Supervisors \( (M = 4.6, SD = .49 \text{ vs. } M = 4.30, SD = .63, p = .18). \)

Most participants felt the training provided them with “the right tool to introduce dialogue and engage the person directly” and “to ask the questions directly,” (Transportation Supervisor #13).

**Perceived Comfort Level when Intervening with a Suicidal Patron**

The vast majority of study participants indicated the training program increased their comfort level when intervening with a suicidal patron. A mean rating of 4.43 (SD = .6) was observed for this item, indicating that the average self-rating of comfort with suicide intervention may be described as between “comfortable” and “very comfortable.” There were no differences in the self-rated comfort level reported by Special Constables and Transportation Supervisors \( (p = .30). \) Nonetheless, one Transportation Supervisor was more measured in his assessment of his comfort level and candidly admitted: “You are never that comfortable. I don’t think you are ever comfortable, I don’t think anyone is ever comfortable” intervening with a suicidal patron. (Transportation Supervisor #8).

One Transportation Supervisor confessed that prior to attending the training he would have called Transit Control and he would have reported his observations and have Transit Control implement an “unsafe platform” procedures, i.e.: train entering the station must reduce their travelling speed. “But to actually approach the person and to actually bring the subject up was quite different story.” (Transportation Supervisor #4).

**Applying Skills to Workplace**

Skill is “an ability that allows a goal to be achieved within some domain with increasing likelihood as a result of practice” (Rosenbaum & Carlson, 2001, p. 454). Thus, skill acquisition can be considered as attaining capabilities by practice, which helps increase the possibility of goal achievement. Fitts (1964) suggested there are three stages of skill acquisition. In the first
stage, the cognitive stage, a considerable amount of cognitive activity is required to encode the skill into a form that will make it possible for the trainee to generate the desirable behaviour. The primary goal for the trainee in this stage is to know what is to be done rather than improve motor pattern. Thus, improvement is observed through verbal mediation. The second stage of skill acquisition, the associative stage, involves a gradual improvement on performance. Errors can be detected and eliminated by practice so that performance level becomes more consistent. At this stage the learner focuses more on “how to do” rather than “what to do.” In Fitts’ last stage, the autonomous stage, the trainee can automatically perform the desired skill.

Since the training, several study participants had the opportunity to apply their newly acquired skills to incidences that involved suicide risk. What they noted was the change in their approach, assessment of the situation, and their intervention. Prior to attending the training they would have “gone to the location of the individual and just stood there and observed” (Transportation Supervisor #5). After the training reported one of the Transportation Supervisors

I change the way that I approach an individual, um I may have not recognized some of the signs before but now, as I am going in, I am assessing the individual, seeing what their mood is, uh seeing their potential for self harm…before I might have overlooked these details (Transportation Supervisor #6).

Go in and just intervene…I found myself actually more aggressive since the training. I take the platform personally now. It’s my station, it ain’t happening here on my shift and not after I go home. (Transportation Supervisor #12).

One Supervisor admitted that prior to the training he would not have dealt with the situation in the same manner

I would not have approached her…I would not have turn around and say ‘do you want to hurt yourself?’ I wouldn’t have talked as openly as I did…I told her I would be more than happy to help her get assistance…to get her pointed in the right direction. (Transportation Supervisor #3).

Several other study participants recounted incidences where they applied the new skills they acquired in the training and the patrons denied being suicidal.

“I felt confident in what transpired,” said one of the Transportation Supervisors and admitted that prior to the course he would not have
introduced the idea of suicide…I would have talked and tried to find out what their intentions were…what their story was…but I would not asked about suicide. (Transportation Supervisor #13).

One of the Transportation Supervisors on several occasions was physically assaulted by aggressive, mentally ill patrons; however, he emphasized, prior to the training his reluctance to approach a patron who exhibited symptoms of mental illness or distress was not due to “the fear of being assaulted, it was due to the delay in my work”, which he feared most. This concern over being delayed by a situation that might require more of his time than he could afford, he stated, inhibited him from getting involved and intervening. After attending the training, he reported “I don’t look at the delay part.” What matters to him is “saving a human’s life,” he stated. (Transportation Supervisor #4).

**Recommendations for Improving the Suicide Prevention Program:**

Study participants were asked to offer suggestions on ways to improve the suicide prevention training. Most participants suggested instituting a refresher course for their training to be up-to-date. As one study participant succinctly said “I don’t think it’s something we should just say ‘okay we have done it’ I think it’s important to keep follow-ups,” (Transportation Supervisor #6) and as another Transportation Supervisor admitted:

> You always tend to forget things that you were taught that you don’t review, or don’t use…what’s hanging in the balance here is the most paramount issue…people’s lives. So when you have something of such great concern, you have to pay more attention to it. (Special Constable #21).

There was no agreement among study participants whether the refresher training course should be offered annually or biannually. Some study participants suggested the refresher training to maintain skills ought to be “offered on a yearly basis, once a year for all employees, all frontline employees” (Special Constable #25) “for civilian people who aren’t dealing with it daily, it’s like a CPR.” (Special Constable #29), and should be “at least four hours long” (Transportation Supervisor #11). The need for refresher training is paramount in the opinion of one Transportation Supervisor: “I need to retain knowledge. I have to retain it, it’s a matter of survival...to be able to deal with it,” he explained. (Transportation Supervisor #2).
Some study participants had suggested the suicide prevention workshop should be “attached to the recertification training” of those employees whose occupational roles within the TTC required recertification training on a regular basis, i.e.: Train Operators, Collectors, maintenance workers. It was also suggested that the suicide prevention training needs to be “mandatory for all new recruits” (Transportation Supervisor #11) and incorporated in their training.

While some study participants thought that “updates in the format of letters, e-mails would be sufficient” (Transportation Supervisors #3, #12, #6) one of the Special Constables candidly admitted: “Most people don’t read written material.” (Special Constable #18). There were several aspects of the training study participants wished to see improved such as location of workshop delivery and content of the training.

To improve the suicide prevention program there was a need to “raise awareness by training more TTC employees. The Training program should not be centralized in one location, it should be taken to different divisions...make it an open forum employees can sign up for the program and do it on a rotating basis.” (Transportation Supervisor #2). It was also recommended the training for Transportation Supervisors consist of a module describing the Mental Health Act and explaining role and legal powers of the Special Constables. To understand and be more aware one Transportation Supervisor (#12) suggested the refresher training should include “guest speakers, people who have attempted suicide” in the subway to explain why they chose “such violent way to do it.” (Transportation Supervisor #12).

Most Special Constables recommended separate “more in-depth, almost a professional level suicide prevention program for the Special Constables” (Special Constable #24) that focuses on skills enhancement as “many of us have a lot of skills that we could work on” stated one of the Special Constables. (Special Constable #24). The refresher training one Special Constable suggested should also include a module reviewing Special Constables’ responsibilities. For new recruits to the force it was suggested it was essential to “train those young officers on how to go on the platform in a crowded subway station because lots of time you’re dealing with a dangerous person in a dangerous environment and how to actually take care of it. (Special Constable #19).

It was also suggested to integrate into the training a component that guides young officers through scenarios of the different types of mentally ill people and different states of suicidality and actually train those young officers on how to go on the platform in a crowded subway
station… Role playing scenarios because you know there’s more to it than simply dealing with the mentally ill or suicidal person who walks around the subway platform with trains going by, with hundreds of people around. You have to be very careful. (Special Constable #24).

One of the Transportation Supervisors (#7) suggested the suicide prevention training workshop should be taken out of the classroom and into “Bay lower station” which is no longer in use and can accommodate a simulated suicide. He felt it would provide hands-on experience in a “real environment” for the trainees.

**Summary**

Several themes emerged from the study participants’ reactions to the *safeTALK* training. Consistent with Alliger et al. (1997) these emergent themes can be separated into two categories: affective and utility reactions. Participants’ reactions to the content and format of the *safeTALK* workshop, the skill training methods, and perceptions of trainers’ performance can be categorized as affective reactions while the feedback regarding the usefulness and relevance of the training to their work is consistent with utility reactions.

In both these categories there were mixed reactions to the training from Special Constables and Transportation Supervisors. Overall, Transportation Supervisors reacted more favourably to the training than Special Constables. Most Transportation Supervisors indicated the workshops had positive effects on their knowledge and their capacity to intervene with a suicidal patron. Conversely, Special Constables stated the training was too basic for their training needs and their occupational responsibilities and legal powers. While most interviewed Transportation Supervisors did not identify their acquired new skills as intervention skills, their descriptions of their new acquired abilities are consistent with the acquisition of intervention skills.

All of the study participants indicated they were competent to intervene and were confident and comfortable interacting with distressed patrons. The training dispelled some of the myths about suicide and assuaged their fears of communicating directly and openly about suicide intent. Subsequently, they were more willing and keen to intervene with patrons in situations of suicide risk. While participants felt the training did not influence their attitudes toward suicide, they vehemently stated the training raised their suicide awareness and increased their vigilance when patrolling the subway. The comments and observations made by several of the Special
Constables need consideration when planning future suicide prevention training for Special Constables.
CHAPTER SEVEN
CONCLUSIONS AND DISCUSSION

This study provided a comprehensive evaluation of the immediate and long-term effects of two programs implemented by the Toronto Transit Commission (TTC), safeTALK and suicideAWARE. The study hypotheses focused on training effects on knowledge, attitudes, and intervention abilities within a real world context, evaluating non-clinical employees in the workplace.

The analytical framework, the Kirkpatrick Model (TKM), was concomitantly appraised for its efficacy in evaluating various aspects of the training. This study is the first to test the applicability of the TKM as an analytical framework to guide suicide prevention program evaluation. The current study evaluated the safeTALK and suicideAWARE programs across two levels of the Kirkpatrick Model of evaluation (Kirkpatrick, 1967): reactions (Level I), learning (Level II) and Levels III (behaviour). Level IV (results) of the Kirkpatrick Model were beyond the scope of this study.

This final chapter provides a discussion of the program results and their implications, based on the Kirkpatrick Model; an evaluation of the model as a theoretical framework for gatekeeper programs; the strengths and limitations of the current study; implications and future research considerations; and concluding remarks.

The Study Results and their Implications

The training evaluated was a gatekeeper suicide prevention program which consisted of two suicide prevention programs, safeTALK and suicideAWARE, which were designed by and adapted for the TTC by LivingWorks. The three-hour long suicideAWARE workshop was delivered to Train Operators as a component of their re-certification training; the one-day safeTALK workshop was provided to TTC Special Constables, Transportation Supervisors, Training Instructors, and occupational health and safety personnel. Attendance in the training was mandatory for Transportation Supervisors and TTC Special Constables. For all other TTC personal attendance was voluntary. The objectives of the suicide prevention programs were to sensitize TTC employees to suicide prevention and to increase their ability to identify and refer distressed patrons to appropriate resources.
The evaluation was guided by the following three study hypotheses: 1. Factual knowledge about suicide and its risk factors would increase after the training and would be maintained over time. 2. Positive attitudes towards suicide intervention would increase after training and would be maintained over time. 3. Suicide assessment and intervention skills would improve and be maintained over time.

Program participants were asked a series of questions (through personal interviews) to evaluate their reactions to the training program. The evaluation questions focused specifically on these three aspects of the training objectives: knowledge, attitudes, and intervention skills. The study used a two-phase, sequential mixed-method approach of converging quantitative and qualitative methodologies: a quantitative study of the effect of the gatekeeper program on attitudes, knowledge, and intervention abilities, followed by a qualitative study of TTC Special Constables and Transportation Supervisors satisfaction with the program and its impact on their competence, confidence and willingness to engage and intervene with individuals at risk. The analysis of these findings answered the research hypotheses.

The Kirkpatrick evaluation model was used as a theoretical framework to guide the program evaluation. It was concomitantly appraised for its efficacy in evaluating various aspects of the training as well as its effectiveness as an evaluative tool of gatekeeper suicide prevention programs.

The results of the study supported the three hypotheses and, thus, also supported the research questions. The detailed results for each hypothesis are organized under two major sub-headings corresponding to the two levels of the Kirkpatrick evaluation model evaluated in this study: Level I - Reactions and Level II – Learning. Levels III and Level IV of the evaluation model were not evaluated for this study; the reason for their exclusion is briefly discussed at the end of this section.

**TKM Level I: Reactions to the Training**

Training reactions can be divided into two basic elements, affective and utility reactions (Alliger, Tennenbaum, Bennett, Traver & Shotland, 1997). The participants’ responses to the study interviews noted both affective and utility reactions to the safeTALK program.

The initiative taken by the TTC to implement an educational suicide prevention program for employees was praised by all study participants. The overall reaction of Transportation
Supervisors was positive and the program was considered constructive, beneficial, and relevant to their work. It was felt that the training program contributed to an increase in suicide awareness, dispelled myths about suicide, and enhanced knowledge of intervention skills.

Reactions to the training content, format and teaching strategies were generally positive, with the exception of Special Constables, who felt the training did not meet their specific training needs. To maintain their knowledge and skills, the vast majority of participants recommended that the TTC provide a refresher course on a regular basis.

Anecdotal and other evidence suggests that post-training reaction questionnaires are the most widely used evaluation criteria in applied settings. Reaction criteria which are operationalized by using self-report measures represent trainees’ affective and attitudinal responses to the training program; however, “reactions measures are not suitable surrogates for other indexes of training effectiveness” (Tennenbaum & Yukl, 1992, p. 425). Affective reaction measures rarely show variance, as most trainees react positively to all training experiences, and have been repeatedly shown to be unrelated to other levels of training success such as learning and transfer of acquired knowledge in the workplace (Alliger & Janak, 1989). Thus, a training evaluation that relies solely on affective reaction measurements is an unacceptable assessment of training effectiveness. At the same time, there is some evidence that the role of training reactions in explaining training effectiveness is more complex than has been indicated in previous models of training criteria (Alliger & Janak, 1989). Reactions may also play a complex indirect role on other levels of training criteria, such as learning and behaviour. Utility based reactions measures showed higher levels of correlation with learning or outcome measures than did affective reactions measures (Alliger, Tennenbaum, Bennet, Traver, & Shotland, 1997).

**TKM Level II: Intermediate and Long-term Training Effects on Knowledge (Hypothesis #1)**

The training objectives relating to the acquisition of suicide knowledge were met by the training. These included identifying behavioural cues which might be indicative of suicide risk, organizational procedural protocols, and intervention knowledge. Participants showed significant gains in appropriate procedural responses as measured by the Suicide-risk Procedural Questionnaire (SPQ). The post-training gains were in all areas of suicide knowledge, particularly in the warning signs and appropriate procedural response that need to be initiated when suicide
behaviour is suspected in the subway station. The follow-up at three months indicated that the participants had retained these knowledge gains.

The training had a differential effect on the three occupational groups. Of the three, Special Constables showed the greatest benefit, demonstrating a significant increase in knowledge of warning signs and procedural actions. The proportion of Special Constables with a perfect SPQ score nearly doubled after the training (52 percent versus 100 percent). While Transportation Supervisors also showed significant gains in knowledge, Train Operators showed the least improvement in their knowledge of warning signs and procedures.

At the three-month follow-up, differences were noted in knowledge gain retention among the occupational groups. The number of Special Constables who retained their knowledge gains remained significant compared to pre-training. Train Operators not only retained their gains in knowledge but they continued to show further gains, as was evident from the differences between post-training and follow-up scores. Only Transportation Supervisors showed minimal and insignificant retention and their post-training scores were only slightly higher than those at pre-training.

All safeTALK participants showed significant gains in intervention knowledge specifically as measured by the modified Intervention Knowledge Test (IKT-M). Immediately following the training, a notable increase in knowledge was demonstrated by a greater number of participants, a 36 percent relative increase in participants’ mean scores. At the three-month follow-up, most of the gains in knowledge were retained compared to pre-training scores and the training continued to show a considerable effect on knowledge.

The training did not have differential effects on the Special Constables or Transportation Supervisors as was demonstrated by the lack of a statistically significant interaction between the training and the occupational group; however, Transportation Supervisors showed greater gains in intervention knowledge immediately following the training and at the three-month follow-up.

**TKM Level II: Immediate and Long-term Training Effects on Attitudes**

(Hypothesis #2)

Attitudes towards suicide may serve as predisposing, reinforcing and enabling factors in suicide intervention (Tierney, 1988). The safeTALK and suicideAWARE attitude modules were designed to increase awareness of attitudinal barriers about suicide and foster desirable attitudes
towards suicidal individuals. The training encouraged participants to explore their own attitudes and those of other participants to suicide and suicide prevention. It guided participants to reflect on the impact negative attitudes may have on suicide intervention and on their willingness to intervene with a distressed and potentially suicidal individual. One of the key objectives of the study was to evaluate the effects of the training on participants’ attitudes.

The majority of participants endorsed a greater number of positive attitudes towards suicide and suicide intervention after completing the training. Participant responses recognized that anyone has the potential to be a suicide victim regardless of age, family situation and/or economic status, and that suicide might present itself as a solution to one’s problems at some point in anyone’s life. Participants noted that suicidal behaviour is not impulsive, but is a “cry for help” and that threat of suicide should be taken seriously. Intervening with a person who is suicidal was viewed as the right thing to do, as most suicidal individuals do not sincerely wish to die and can be dissuaded from acting by a “friendly ear.” Participants understood that whereas suicidality is a transitory state of mind, the suicide risk remains once the immediate crisis is over. These attitudes were noted consistently in participants’ personal interviews and responses on the Suicide Opinion questionnaire.

Overall, participants continued to express a comparable pattern of positive attitudes at the three-month follow-up. Positive attitudes towards suicidal behaviour and suicide intervention were maintained by the majority of participants compared to pre-training. These results are consistent with findings of previously evaluated suicide prevention training programs that also found that suicide prevention training had a positive effect on participants understanding of suicide, resulting in greater sensitivity and positive attitudes towards suicide prevention (MacDonald, 1999; Tierney, 1988; 1994).

However, the occupational groups differed in terms of positive attitudes towards suicide prevention, both pre-and post-training. From the outset, the majority of Train Operators endorsed a greater number of positive attitudes. However, improvements in positive attitudes post-training were not homogeneous across all occupational groups, nor were all the observed improvements statistically significant. The training had the least effect on Special Constables and Transportation Supervisors; they continued to endorse negative attitudes towards suicide and suicide prevention.
The differences between the three occupational groups were still evident at the three-month follow-up. Train Operators maintained their positive attitude gains; Special Constables and Transportation Supervisors did not alter their attitudes towards suicide and suicide prevention. Comments made in interviews with Special Constables and Transportation Supervisors suggest that their attitudes towards suicide were deeply ingrained and shaped by personal and cultural beliefs and workplace exposure to suicide and its aftermath. However, the training may have been effective in elucidating participants’ attitudes and beliefs as evident from the decrease in the number of undecided responses across attitudinal items.

The training appears to have provided the most gains to the Train Operators, further enhancing their positive attitudes toward suicide prevention. Although the study findings appear to support the second hypothesis, caution should be exercised when interpreting the immediate and long-term effects of the training on the different occupational groups.

TKM Level II: Immediate and Long-term Effects on Acquisition of Intervention Skills (Hypothesis #3)

To measure whether the suicide intervention abilities of the participants of the safeTALK program were increased by the training, the participants were assessed in their ability to distinguish between facilitative and deleterious hypothetical intervention responses to a suicidal client, indicating an awareness of appropriate suicide intervention skills. Significant gains in intervention skills were observed following the training, with more participants able to identify a greater number of responses correctly. At the three-month follow-up the gains in intervention skills were maintained, as compared to pre-training; additionally it was observed that there was a further increase in intervention skills at the three-month mark.

The training appears to have had differential effects on the two occupational groups i.e., TTC Special Constables and Transportation Supervisors. A comparison of the two occupational groups revealed Transportations Supervisors demonstrated greater gains in intervention skills acquisition immediately following the training as well as greater retention of intervention abilities.

The findings of this study demonstrated the safeTALK and suicideAWARE programs had positive immediate and long-term effects on participants’ knowledge of suicide and suicide prevention, attitudes, and intervention skills; hence, the study hypotheses were confirmed. While
gains in knowledge, enhancement of positive attitudes and intervention skills were demonstrated, it is unknown whether the changes are sustainable for the individual participants beyond the three-month follow-up period, or if there will be any long-term impacts of the training on workplace interactions with TTC patrons.

The quantitative and qualitative findings were triangulated across assessments: knowledge, attitudes and intervention skills and interviews. The quantitative and qualitative data triangulation yield a more complete analysis and added complexity to the evaluation findings and revealed different aspects of the training reality. By “zooming in and zooming out with a lens” (Willems & Raush, 1969, pp. 82-83) it allowed the researcher to focus on different aspects of the training and construct a comprehensive picture of issues related to attitudes toward suicide, conditions, and contexts. While the qualitative evaluation findings provided depth to some of the causes behind changes and offers additional explanations, the quantitative methods provided a yardstick for measuring change. Triangulation of qualitative and quantitative of data permitted elaboration, enhancement, clarification, and illustration of findings. It explored interconnected and/or distinct aspects of the data and increased the breadth and depth of the findings and interpretation by analyzing them from different perspectives.

The quantitative results of this study indicated the gatekeeper training program implemented at the TTC had a proximal and distal medium- to large-size effect on increasing participants’ knowledge about suicidal behaviour and risk factors, and suicide risk assessment and intervention skills. Similarly, the training had a medium-size effect on desirable positive attitudes toward suicide prevention. However, there was substantial occupational group variation in terms of suicide-risk related procedural knowledge with the greatest proximal and distal gains demonstrated by the Special Constables and the smallest proximal gain demonstrated by the Train Operators. At the 3 months follow-up the training effect on Transportation Supervisors’ suicide-risk procedural knowledge was minimal. While the quantitative study demonstrated the training contributed to acquisition of knowledge, enhancement of intervention skills, and desirable attitudes towards suicide, the interviews provided participants’ impressions and appraisal of various aspects of the training and its efficacy in imparting salient knowledge. The qualitative data provided insight into interviewed participants’ affective reactions to the content, format, and teaching strategies and allowed for elaboration and distinctions between affective and utility reactions. Statements made by Special Constables regarding the workshop format and
content clearly conveyed their perception of the training’s relevance to their job-related duties and plainly elucidated the disparities in their perception of training needs in comparison to other occupational groups. The interviews clarify and facilitated the discovery of contradictions in the data. For example, while Special Constables felt that given their previous training and experience, the *safeTALK* workshop was “too basic” for their training needs, the findings from the quantitative measures indicated the training contributed to substantial gains in desirable attitudes toward suicide prevention, knowledge, and intervention skills.

Thus, triangulation of the quantitative and qualitative results captured a more complete, holistic and contextual portrayal of the training, and offered in-depth understanding of important and germane aspects of the training and its effects on participants’ knowledge and intervention skills. Examining the training from multiple perspectives enhanced our understanding by allowing for new and deeper dimensions of the training experience to emerge.

Several aspects of the training have been identified by Special Constables and Transportation Supervisors as salient aspects of knowledge acquisition: 1. Warning signs- participants were taught that suicide warning signs are at once observable behaviours, pre-suicide indicators and danger alarms that provide the unique, and sometimes the only, opportunity to intervene in a developing suicide crisis. The list of behavioural warning signs was repeatedly identified by Transportation Supervisors as novel valuable knowledge which was considered beneficial in the identification of distressed at-risk for suicide patrons. 2. Intervention skills- Prior to attending the training, Special Constables and Transportation Supervisors reported that observing behaviours that were indicative of distress and potential suicide risk engendered trepidation and a misplaced sense of responsibility for the potential possible actions of suicidal patrons. Although they felt compassion, they were reluctant and fearful of asking a distressed patron whether they were suicidal. Experiencing strong emotional reactions in a possible suicide crisis may have inhibited a helpful response. Being taught to engage distressed patrons in direct and open discussion regarding their suicide intent was cited by Special Constables and Transportation Supervisors as the most important intervention skill learned in the gatekeeper training program.

Triangulation of the data demonstrated that using a mixed method approach did not result in a single, clear-cut, consistent picture, but rather presented a challenge to improve comprehension of the various reasons for the existence of inconsistencies between the two sets
of data. The following illustrates such divergent findings and exemplifies the contradictions between participants’ perceptions of the training’s effect on their attitudes and the quantitative results.

One of the key modules in the safeTALK workshop was a group discussion of attitudes to suicide. The purpose of the group discussion was to help participants explore their own attitudes and those of other participants toward suicide and suicide prevention and to reflect on how their attitudes might affect their willingness to intervene, and the intervention process. However, most Special Constables and Transportation Supervisors indicated the training did not influence their attitudes toward suicide; nor did they feel their attitudes toward suicide were altered as a result of the training. They attributed their attitudes toward suicide to be deep-rooted in personal religious and/or cultural beliefs, their familial experience with mental illness and suicide, and their work experience with the clean-ups in the aftermath of suicides in the subway system. They felt their attitudes were intrinsic to their beliefs and defined who they were, hence attending the suicide prevention program was not sufficient to alter their attitudes nor did they feel the training challenged attitudes toward suicide and suicide prevention. The training, however, helped them enhance their understanding of the suicidal patron. Examining the proximal findings of the Suicide Opinion questionnaire indicates the training had a small-size effect on Special Constables and Transportation Supervisors positive attitudes towards suicide prevention. However, there was a substantial occupational group variation. The proximal effect size of the training on positive attitudes of Train Operators was large. Furthermore, at pre, post, and follow-up Train Operators endorsed a greater number of positive attitudes toward suicide prevention than Special Constables or Transportation Supervisors. Conversely, at the 3-month follow-up Special Constables and Transportation Supervisors endorsed a greater number of positive attitudes and thus the training had a medium-size effect on their desirable positive attitudes.

Sechrest and Scott (1993) asserted that personal characteristics can not be influenced by training; however, certain negative attitudes can be modified with adequate information, even towards an emotionally, culturally, and socially charged issue as suicide. There is an association between knowledge and personal characteristics: appropriate conduct in an intervention is determined by the presence or absence of certain personal characteristics, or interpersonal abilities rather than theoretical knowledge (Sechrest & Scott, 1993). Additionally, knowledge
acquisition is limited by certain personal and interpersonal characteristics and communication, listening, and empathy abilities (Whiteman, Hartman & Brannon, 1982). This illuminates the importance of certain intrinsic qualities which are essential for acquisition and/or the development of specific competencies. However, personal characteristics such as attitudes influenced by rigid personality traits are not modifiable (Sechrest & Scott, 1993). For example, believing that those who threaten or attempt suicide are attention seekers or that those who threaten suicide never do suicide will influence the intervention or their willingness to intervene. Thus, in addition to knowledge acquisition, certain attitudes can be a barrier to intervening with or adopting an empathetic attitude toward the suicidal person (Tierney, 1988). McIntosh (1985) argued that knowledge acquisition can have a positive effect on attitudes. The dissemination of precise information on suicide can contribute to the development of positive attitudes and does allow to adequately identifying a suicidal individual and the available resources to assist them. Thus, a suicide prevention training program can sensitize and influence individuals to a higher degree of personal responsibility and willingness to intervene (Tierney, 1988; 1994).

Learning gains are expected to lead to the transfer of newly acquired knowledge and skills to the individual behaviour in the workplace (Level III of the Model). Sechrest and Scott (1993) argue that knowledge acquisition is essential in skill training as it allows for better understanding for and the impact actions have. Where suicidal individuals are concerned, it has been demonstrated that knowledge facilitates the identification of individuals at risk for suicide (Tierney, 1988; 1994). Additionally, there is an association between specific knowledge and intervention skills; however, the association between knowledge and skills is not always clear in practice (Tierney, 1988; 1994).

TKM Level III: Behaviour

Following the Kirkpatrick Model, learning is measured during training and refers to attitudinal, cognitive, and behavioural learning. Behaviour refers to the application of the acquired knowledge and skills to on-the-job performance. Evaluating at this level attempts to answer the question: Are the newly acquired skills, knowledge or attitude being transferred to the workplace?

Because the current study design did not include direct measure of the transfer of acquired knowledge and skills to the workplace setting, evidence for this third level is based
exclusively on interviews with the participants post-training and represents participants’ perceptions of their behavioural changes. According to these reports, the training increased participants’ alertness to the warning signs of patrons’ distress, as well as increased their vigilance, confidence, comfort and overall sense of perceived competence and confidence to intervene with a suicidal patron. They also noted changes in their approach, risk assessment and intervention skills.

**TKM Level IV: Results**

The present research did not set out to evaluate at the level IV of the Kirkpatrick Model, i.e. training effects on organizational results. Suicide rates reflect a multitude of “complex interactions, of risk and protective factors that operate outside the suicide prevention program influence” (Kalevald & English, 2005).

The ideal demonstration of the effectiveness of a suicide prevention program would be the direct reduction of suicide rates. However, there are substantial difficulties in demonstrating the impact of training on suicide rates. Suicide rates vary in unpredictable ways, with time, geographical regions, age groups, or special populations. “Local suicide rates, due to significant fluctuations that occur in small populations, are often not useful in evaluating effectiveness of suicide prevention programs.” (Kalevald & English, 2005:132). Suicide is a statistically rare event and the effect of training may not be evident for many years because of the limited sample. Additionally, there are no indicators to inform us as to how many employees need to be trained and what level of contact they need to have with patrons at risk in order to impact suicide rates. Moreover, suicide rates are affected by a multitude of societal and individual factors, including suicide prevention programs (Beautrais, 1998). Given these complexities, it would be speculative to attribute any changes to a single, specific intervention.

**An Evaluation of the Kirkpatrick Model as a Theoretical Framework for Gatekeeper Programs**

Recently, an international expert consensus report called for increased empirical investigation to further develop the evidence base for suicide prevention research (Mann et al., 2005). The same expert panel upon reviewing the evidence found that as many as 80% of those who died by suicide had untreated mental illness (Mann et al., 2005). These results illuminate
the need for increased efforts within the field of community-based approaches to suicide prevention. Several studies (Mann et al., 2005; Goldsmith, 2002) have identified gatekeeper training as a promising strategy for suicide prevention.

As a universal and possible selective strategy, gatekeeper training is designed to train those in strategic positions in early identification and detection of individuals at high risk for suicide and to improve appropriate and timely referrals to mental health resources in the community. The efficacy of gatekeeper programs is supported by theoretical rationale (Goldsmith et al., 2002, Mann et al., 2005) and has an endorsement from the Center for Disease Control (1994). More recently, a review of thirteen gatekeeper training studies from seven countries by Isaac, Elias, Katz, Belik, Deane, Enns and Sareen (2009) found gatekeeper training to be effective at imparting knowledge, teaching intervention skills, and enhancing positive attitudes toward the suicidal individual and suicide prevention. Moreover, six of the thirteen studies demonstrated significant decrease in the number of suicide attempts and deaths by suicide.

There is a widespread need to improve the evaluation of suicide prevention programs and it is pertinent that we learn and understand more about the factors predictive of outcomes at different levels of evaluation. Diverse factors impact outcomes. The diversity of these factors stems from the fact that suicide prevention occurs at multiple levels and within diverse settings within a workplace or community. Suicide prevention programs do not always consider the diversity of sub-groups and/or social context. Nor are suicide prevention programs at all times relevant to those individuals who are being targeted as trainees, and moreover, interventions skills being taught are not always appropriate for those being trained. Gatekeeper suicide prevention program evaluation needs to be multi-stage process involving the setting of objectives, the execution of strategies, the collection of data, and an assessment or appraisal of relative success or failure of a the program. Glasgow, Voget and Boles (1999) proposed a comprehensive evaluation framework, and have argued that multifaceted interventions should be evaluated with measurements suited to their settings, goals, and purpose. Kraiger (2002) emphasized the multidimensional target areas of evaluation: training content and design i.e., design, delivery, and validity of training, changes in learner i.e., affective, cognitive, and behavioural, and organizational payoff i.e., transfer climate, job performance, and results.
To date the vast majority of suicide prevention programs assume efficacy with little scientific evidence. Data about the ultimate program success or failures are extremely limited. Program development in this area is generally motivated by belief and ideology rather than evidence about effectiveness. Outcomes must be based on evidence i.e., being able to make a case for its effectiveness and contribution to suicide prevention. Therefore, the most important evaluation question for suicide prevention program is, “Does the training activity of the program makes it more likely that individuals in the community will follow through and intervene and thus prevent suicide or self harm?” Collecting scientifically reliable evidence for effectiveness is not always possible, especially for prevention programs. How can we know if any instances of suicide have been prevented by this program? What do we measure in order to know that suicide has been prevented by this particular program?

When evaluating suicide prevention program several aspects of evaluation need to be address: methodological (what, how, and when to measure), substantive (how training should be evaluated), and organizational concern (what role does the organization play in training effectiveness).

Introduced in 1959, the Kirkpatrick Model (TKM) has stood the test of critical review, gaining support over time to be one of the most widely accepted and influential models of program evaluation (Phillips, 1996; 2003). Most of the evaluation models found in the literature are generally based upon Kirkpatrick’s four levels (Bomberger, 2003; DeSimone & Harris, 2002). Despite criticism and the development of other comprehensive evaluation models, TKM is still widely utilized, due to its simplicity and practicality (Kirkpatrick & Kirkpatrick, 2006; Phillips, 2003). The Model’s strengths and popularity can be attributed to its potential for simplifying the multifaceted and complex process of training evaluation. Moreover, the framework provides a rigorous approach to conducting evaluations.

It is evident from the suicidology literature that there is a paucity of empirical research evaluating suicide prevention programs. The need for comprehensive, systematic evaluations is well documented (CASP, 2004; 2004; Goldeny, 1998; Tierney, 1994; CDC, 1994; National Task Force, 1987). Lack of sound evaluation methodologies remains one of the most significant barriers to the identification and implementation of effective intervention strategies (Mann et al, 2005; Breton et al 2002; CASP, 2004; Goldeny, 1998; Tierney, 1994; CDC, 1994; National Task Force, 1987).
Barriers to the implementation of summative evaluations involve not only the complexities inherent in the area of suicide prevention but also include cost, limited time and resources, inadequate sample size, and uneven commitment to the evaluation process (Kalevald & English, 2005). Additionally, erroneous assumptions conclude that some training cannot be quantitatively measured, that there are too many variables affecting the behaviour change other than training, and that evaluating training programs is very expensive (Swanson, 2001). Because of these assumptions, suicide prevention program evaluations have long been focused on trainees’ reactions to the training program i.e. Kirkpatrick’s Level I (Bailey, Carpenter, Dickson & Rogers, 2002; Alliger & Janak, 1989).

One of the greatest challenges in suicidology is developing and using evaluation methods. Due to the common misunderstanding of time constraints, costs and the complexity of higher level evaluations, many gatekeeper program evaluation efforts still emphasize the lower level (reactions) of evaluation of the Kirkpatrick Model (Bailey, Carpenter, Dickson & Rogers, 2002). Nonetheless, the Kirkpatrick four level evaluation model still serves as an effective guide for conducting training evaluation even though it has been five decades since its introduction. As stressed by Kirkpatrick (1959; Kirkpatrick & Kirkpatrick, 2006), evaluating the behaviour is more complicated, difficult, and time-consuming than evaluating reactions to the training program and what was learned (Level I and II). Consequently, Kirkpatrick believed Level III is often overlooked (Kirkpatrick & Kirkpatrick, 2006; Kirkpatrick, 1998). Much time, energy, and expense is given to Levels I and II because evaluators have the most control over what is evaluated at these levels.

One of the main objectives of this study was to demonstrate the applicability of the Kirkpatrick evaluation model to gatekeeper program evaluations. However, this study evaluated only Levels I and II of the Model.

**Evaluating the Applicability of the Analytical Framework**

The Kirkpatrick Model forms a logical framework to examine the results and impacts on individual and organizational perspectives (Setaro, 2001). As a conceptual framework, the Model aids in determining what data are to be collected, and serves as a guide to the type of questions that should be asked and the criteria that might be appropriate. The simple pragmatic
model allows various aspects of the training to be evaluated; the four levels answer different questions regarding training objectives and goals.

An important contribution of the framework is how information obtained from the four levels can be integrated to provide an accurate and thorough evaluation of the training itself. Results from Level I (reactions) and Level II (learning) evaluations can be combined with those from Level III (on the job behaviour) to gain a better understanding of the outcomes of the training. In this way, information can be synthesized regarding the extent to which the trainees perceived the training useful, learned the relevant intervention skills, transferred this learning to improve interventions in the workplace, and thereby contributed to the reduction in suicide incidents.

Since the Model focuses the evaluation process on four levels of outcome data, the TKM reduces the measurement demands for training evaluation. Generally, the data is collected after the training has been completed and thus it eliminates the need for—or at least implies that—pre-training measures of learning are not essential for determining program effectiveness (Bates, 2004).

The Kirkpatrick Model is outcome and objective-oriented and focuses on determining the effectiveness of a program. The evaluation model provides the ability to capture the trainees’ progression from their initial reactions to changes in their behaviour as a result of the training. In other words, it is a summative evaluation model: the evaluation is conducted only after the training has occurred, in order to assess the merit and worth of the training program, and provide a summary report of the training outcomes. The evaluation at each level answers whether a fundamental requirement of the training program was met. Based on the evaluation results, Kirkpatrick noted, decisions to continue or alter the training program can be made accordingly: the summative evaluation can be used as a formative evaluation to measure development, future program improvements, and/or other program modifications (Clark, 2008; Kirkpatrick & Kirkpatrick, 2006). In this way, the Kirkpatrick Model offers more than just an analytical framework for program evaluation; it can also be utilized as planning tool for program planning (Clark, 2008).

The findings from this study indicate that learning occurred (Level II) and trainees’ perception of their intervention performance improved (Level III) as a result of attending the
gatekeeper training. Interviewed trainees attributed the changes in their alertness, vigilance and intervention competencies to the knowledge they acquired from the training.

The study provides the groundwork for additional research into the effectiveness of the Kirkpatrick evaluation model as a whole, and also evaluations at each level:

Level I: There is an implicit assumption within the Kirkpatrick Model that each level of training outcome affects the next level in the hierarchy, thus participants’ satisfaction influences their propensity to study, which, if it becomes real learning, can modify behaviour. While a positive reaction may impact learning on its own accord it does not guarantee that significant learning occurred. Alternatively, a negative reaction to the training in all likelihood reduces the possibility that behaviour will change.

Although probing whether participants were satisfied with the training they received does not provide an in-depth understanding of the effectiveness of the training; nonetheless, an examination of participants’ reactions to their training, and understanding the factors that affect participants’ reactions contributes to more effective planning, design and evaluation of suicide prevention programs.

Reactions are emotionally-based opinions. Alliger, Tennebaum, Bennett, Traver & Shotland (1997) divided training reactions into two basic components, affective and utility reactions. They defined affective reactions as enjoyment and satisfaction with the training, while utility reactions referred to perceived relevance and usefulness of the training for subsequent job performance.

Reaction evaluation gauges the interest, and motivation of the participants to the program. Trainee reaction evaluations provide a basis for developing a balanced set of measures, as long as data are provided that can improve facilitation and program implementation. Kirkpatrick emphasizes that Level I is important because positive training experiences may well have beneficial impact on participants’ attitudes and behaviour (Rhoades & Eisenberger, 2002) including their motivation and willingness to engage in organizational commitment and attend future programs i.e.: suicide prevention.

In contrast, negative comments about the program may discourage trainees from attending or completing the program. It is also a key source of information on how to improve future training program (Kirkpatrick, 1998, p.17). In the present study, reactions of participants
to the gatekeeper suicide prevention program implemented at the TTC were positive and provided recommendations for program improvement.

Level II: Learning evaluation examines whether trainees’ attitudes have changed, knowledge improved, and/or skills increased as a result of attending the training program (Kirkpatrick, 1959; 1998; Kirkpatrick & Kirkpatrick, 2006). Learning refers to the cognitive process of acquisition of new knowledge, modification in existing knowledge, behaviour potential, skills, values, and/or understanding, that occurs as a result of reinforced practice, study, instruction, and/or experience (Mikulas, 1977).

Evaluations at Level II focus on the acquisition of declarative or procedural knowledge acquisition, skill development and enhancement, and/or attitudinal change. Declarative knowledge refers to factual knowledge and information that a person knows, while procedural knowledge refers to the knowledge exercised in the performance of some task or activity.

All skill acquisition starts out as declarative knowledge while procedural knowledge is acquired through inferences from already existing knowledge. According to Anderson (1975; 1983), learning declarative knowledge helps set the stage for procedural knowledge. Declarative knowledge has to be present to form procedural knowledge.

Affective outcomes include changes in attitude, motivation and goals that are relative to the objectives of the training workshop. These outcomes are analogous to Kirkpatrick’s learning outcome level (Kraiger, Ford & Salas, 1993).

It is evident from the suicide prevention evaluation literature that Level II evaluations are still one of the most popular forms to evaluate training program effectiveness, despite research that does not support that acquired knowledge and skills equates to behavioural changes in the workplace (Bersin, 2003; Strunk, 1999). However, Kirkpatrick stressed that evaluating learning is important to validating changes in behaviour and determining whether learning is transferable to the workplace. In this study, the proximal and distal effects of the training on TTC employees’ scores was used to for measuring Level II learning performance. Level III: Trainee behaviour/performance is measured by determining the extent to which trainees apply their newly acquired knowledge and skills in the workplace. This level is critical, as it addresses the issue of learning transfer. If trainees cannot apply what they learned to the workplace and demonstrate a positive and measurable change in behaviour, the training effort cannot have an
impact on the organizational results (Level IV). In this study, the identified workplace performance variables were alertness, vigilance, intervention skills, confidence and competence to intervene. The research findings indicated that trainees perceived their intervention competencies improved significantly after the training. However, the study did not directly measure transfer of newly acquired knowledge and skills to the workplace settings. Nonetheless, trainees’ comments on changes in their intervention approaches and behaviour in the workplace demonstrated, albeit indirectly, performance improvement and established the link between Level II and III.

Level IV: Measuring results is the most important and also the most challenging aspect of the evaluation process, yet it is critical for programs designed to influence impact measures. (Kirkpatrick, 1960, Phillips, 1996). According to the literature (Phillips, 1996) the most significant barrier to evaluation at this level is the cost associated with its execution; however, the attempt is essential as Level IV results are often used to justify whether to continue or discontinue training programs (Phillips, 1996).

While the present study did not directly measure the impact of the training on subway suicide rates, some preliminary analyses were conducted from data obtained from the TTC (V. Cosentino, personal communication, January 5, 2011). In the eight years prior to the safeTALK and suicideAWARE training programs (January 1998 to December 2005) there were a total of 129 suicides in the subway system with the number of suicides per year ranging from 12 to 22 with a mean of 16 suicides per year. In the two years following the suicide prevention training (January 2006 to December 2007) the number of suicides ranged from 8 to 13, and the mean number of suicide for each of those years was 10.5, a 35 percent decrease in the mean number of suicides. Conversely, there was no change in the number of suicide attempts following the training.

Critique and Problematic Assumptions of TKM

The Kirkpatrick Model suggests the existence of a hierarchy of ascendancy within the model and implies a strong causal correlation between the four levels. However, research over the past five decades has shown there is little correlation between the participant’s reaction, the training, and learning, and even less between reaction and performance (Alliger et al., 1997; Ruona et al., 2002). In general, the evaluation data generated by the lower levels of the Model do
not buttress the conclusion that the training caused changes in outcomes at the result level. Any number of factors can contribute and confound changes at the results level. Most training efforts have little capacity to directly affect Level IV criteria as the majority of training programs in organizations are short in duration (1 or 2 days) and are meant to have only circumscribed impacts on the training participants.

Exploring some of the difficulties of the Kirkpatrick Model does not represent an argument against utilizing it as an analytical framework for gatekeeper suicide prevention program evaluation. The purpose of examining some of the Model’s limitations is to provide a balanced view of the Model’s strengths and acknowledge some of its weaknesses. Further research is recommended to examine the sequential relationships among the four evaluation levels and their impact on the results.

The findings from this study support the applicability of TKM as a guiding analytical framework for the evaluation of gatekeeper suicide prevention programs. As shown there are several limitations to the Kirkpatrick model that have to be considered before applying the Kirkpatrick model as analytical framework to gatekeeper program evaluation. Kirkpatrick’s model is not readily applicable to gatekeeper’s evaluation without a careful re-evaluation of the model. Such re-evaluation needs to consider the objectives of gatekeeper’s program and outcomes that need to be assessed, the feasibility of carrying out a comprehensive evaluation given time and financial constrains of training programs, and the difficulties with methodologies (appropriate and valid assessment strategies and tools).

However, the framework could be applied as a reference tool to help design high quality suicide prevention program evaluations. Nonetheless, suicidology needs to identify and set unique and critical standards for levels of knowledge and gatekeeper competencies. Because suicide is a multifaceted complex behaviour, a single measure of suicide prevention training will not provide a comprehensive evaluation of the program.

**Strengths and Limitations of the Current Study**

This study offers several important contributions to the research in suicide prevention program evaluations. One of the major strengths of this study is the research design, a sequential mixed - method methodology. The primary advantage of this design is that it allows for a convergence of quantitative and qualitative methodologies, which improves the validity and
credibility of findings and strengthens its scientific rigour. The two training programs, *safeTALK* and *suicideAWARE*, offered by LivingWorks in countries around the world, had not been previously evaluated with such a comprehensive methodology.

Implementing an evaluation of a suicide prevention program in a real-world, non-clinical setting is a strength of this research. The collaboration of academic and workplace resources yielded content-rich and applicable material that can be of benefit to future academic studies as well as other organizations considering suicide prevention programs for their workplace.

The study’s assessment of the efficacy and applicability of the Kirkpatrick Model also contributes to the academic literature. The Kirkpatrick Model has been used in several suicide prevention program evaluations; however, to date, the Model has not been appraised for efficacy in assessing various training aspects of suicide prevention programs. Information drawn from this research, particularly with respect to feedback from the trainees, could be useful to improving program design and delivery and reinforcing the need for both training and refresher courses. Moreover, this contribution to suicide prevention evaluation research and gatekeeper program evaluation research can serve as a basis for future program evaluation strategic planning. Finally, the study validates the need to sustain suicide awareness and prevention programs.

As with any research, in addition to its strengths there are also a number of methodological limitations. One such limitation is the possible selection bias of participants. The participant sample was comprised of trainees from the TTC, with data obtained only from those who self-selected to participate. Nonetheless, the large sample size and diversity of the sample in terms of age, years of experience, occupations (TTC Special Constables, Transportation Supervisors, and Train Operators) is noteworthy. In the present study no specific needs assessment was conducted and there was an assumption that one type of training program would meet the needs of all employees who attended the training, the Transportation Supervisors, Special Constables and Train Operators. Needs analysis takes into account the individual differences of trainees, the organizational climate, objectives, and the characteristics of the tasks to be learned. This information is then used to determine both the method and content of training. In review of training research literature, training needs analysis is recognized as one of the first important “before” contributions to training effectiveness (Salas & Cannon-Bowers, 2001). While the findings in the current study suggest that TTC Special Constables benefited from their training...
and acquired knowledge and skills, most of those Special Constables interviewed articulated the need for more specialized training that reflected their legal powers, their previous training, and years of experience on the job. The lack of a control group in this study is another limitation, suggesting that the results cannot be definitively attributed to the training program. Ideally, the study would be a randomized controlled trial (RCT); however, a very large sample size would be required to demonstrate an impact on suicide in an RCT of training, which runs the risk of contamination between groups. Cohen’s d (Cohen, 1965; 1977) is often considered a crude estimate of effect size and should be interpreted with caution (McGrath & Meyer, 2006); however, in the present study, it provided a useful benchmark for detecting differences from pre- to post-training. The follow-up data was gathered through a mail-out procedure in order to reduce costs and provide a convenient context for participants to complete the extensive package of questionnaires. Data gathered through mail-out may have yielded information that is different than that which was gathered pre-and post-training. The further gains observed on the SPQ and SIRI at the 3 months follow-up period could be attributed to sample attrition due to the low response rate of mailed self-administered questionnaires. The collection of longitudinal data lends itself to the likelihood that some study participants will drop out of the study before all data are collected. Attrition occurs when participants are permanently lost to follow-up. Biases stemming from attrition can lead to overestimation of the effectiveness of the training (i.e., influencing means, variance and correlations among variables).

Kaskowitz and Friendly (1982) pointed out that attrition has two effects on the estimation of training impact. First, as the number of participants used to calculate training gains is reduced, the precision of the gain estimate is reduced as well. The other potential effect of attrition on a training evaluation is the introduction of bias into the gain estimate. In general, the amount of bias will depend on two factors: the amount of attrition and the difference between the gains made by the participants who remain in the evaluation and the gains made by those participants who dropped out. If the attrition rate is high and a large difference in gain exists between the two groups, the bias due to attrition will be substantial.

Furthermore, systematic attrition of the original sample represents a potential threat of bias when those who drop out of the study are systematically different from those who remain in the study. The final sample may form a subgroup that inadequately represents the study participants. The result is that the remaining sample becomes different from the original sample,
resulting in what is known as attrition bias and threatens external validity of the study (i.e., the
generalizability of the findings to other settings or to other individuals). For example, the over-
representation of participants who feel strongly about suicide prevention in the second (post-
training) or third wave (follow-up) of a study may have altered the association between the
training and knowledge acquisition that is different from the true associations in the original
sample and may have resulted in the further observed gains in knowledge and intervention skills.

A comparison of the two groups’ data, to determine whether participants who dropped
out from the study differed systematically from those who remained, was impossible to perform.
The majority of study participants did not provide their TTC employee number and thus the
completed pre, post, and follow-up questionnaires could not be linked as complete cases.

The further gains observed on the SPQ and SIRI at the 3 months follow-up period could
possibly be attributed to participants taking more time to reflect on the questions on the follow-
up research questionnaires. The post-training questionnaires were completed at the end of the
workshop, after a full day of training, when most participants were eager to complete the
research questionnaire and go home. Conversely, the follow-up questionnaires were completed
at home, at leisure, without the added stress of time constraints.

In lieu of the training program being responsible for the further gains at the follow-up, an
alternative explanation to the further increase in knowledge and intervention skills observed at
the 3 months follow-up could be offered by practice effects. The further gains observed may be
an artifact of the participants being retested using the same instrumentation (Gall, Gall, & Borg,
2003) at three temporally close time points. This may have impacted participants’ performance
and observed as further improvement in their knowledge.

Measurements validity is also lacking in this study. Due to time constraints the
measurements used were modified, and thus the psychometric properties of the measures were
not established. As well, using hypothetical statements to evaluate intervention skills represents a
response that is removed from actual intervention experience. Future research methodologies
should strive to evaluate actual suicide intervention capabilities as well as aim to address
questions of measurement validity. Finally, these data are based on a training program designed
specifically adapted for the TTC and delivered to a diverse group of TTC employees. The
characteristics of the sample limit the generalizability of these findings to other workplace
populations, service settings and to other training methods.
Implications and Future Research Considerations

The experience of conducting a comprehensive evaluation of *safeTALK* and *suicideAWARE* using this particular approach and methodology has led to a number of conclusions which have implications for further evaluation. The evaluation designed and implemented in this research included both qualitative and quantitative measurements. The study considered participants’ reaction to the training, in terms of overall satisfaction with the training program and its impact on behaviour. This qualitative data was obtained through personal interviews with trainees. The quantitative aspect of the study measured, through questionnaires, the immediate and long-term training effects of the program on attitudes, knowledge, and intervention skills. Both approaches yielded results which could be applied to the evaluation questions.

Engaging a suicidal individual in helpful dialogue leading to an appropriate referral to professional help/care requires a combination of knowledge, personal confidence in intervention, and demonstrable skills. While suicide risk assessment is a key competency required by a gatekeeper when encountering a potentially suicidal individual (Doyle, 2003; Schwartz & Rogers, 2004) suicide prevention training programs need to offer the opportunity to develop competencies that enable trainees to identify, assess and intervene with individuals in distress and at risk for suicide.

It is essential for suicide prevention programs to develop a “universal of competencies” (Rycus & Hughes, 2000) i.e. a comprehensive inventory of salient competencies which were demonstrated to occur concomitantly in effective intervention. In conjunction with core competencies, suicide prevention training programs need to design and develop training activities which are relevant and particular to each trainee group and are directed to their specific occupational needs.

Future suicide prevention program evaluations need to replace ill-defined learning goals with rigorous, clear outcomes and objectives as a way of assessing the trainee’s learning and competency to intervene with potentially suicidal individuals. They need to describe indicators that reflect increasing levels of competency, and to create methods of evaluation that ensure these indicators have been demonstrated by the trainees according to defined expectations.
Gatekeeper suicide prevention programs designed to improve knowledge and intervention skills lack reliable and valid tools for evaluating their educational efficacy. Suicide prevention program evaluations tend to measure knowledge and skills independently of each other, and assume that the sum of scores from an individual test reflects an accurate representation of the trainee’s competency to intervene. This approach to competency evaluation is flawed. Appropriately robust evaluative measurement tools designed to simultaneously evaluate core gatekeeper competencies need to be developed (Klemp, 1979).

Further refinements of the methodology might also be considered. It would be of interest to study the interrelationships of the training effects among attitudes, knowledge and skills. These were examined as independent variables in this study. The relationship of the SIRI-R to actual suicide intervention dynamics needs to be evaluated. Individual self-reports regarding participant perception of their intervention abilities allows for a less direct, but nonetheless important perspective on the evaluation of intervention abilities. Both the mail-out questionnaires and the interviews proved to be effective methods to gather follow-up data: the questionnaire format is most efficient in providing accurate information at a lower cost; the interviews provided somewhat richer data, and allowed for greater insight into participant’s responses. The questionnaire format allowed for broad sampling; however, it might be shortened as some questions did not provide useful data. Such reduction in the number of questions would make using the questionnaire more efficient.

The workplace is a viable site to train community gatekeepers where employee in-service trainings are a standard component of workforce infrastructure to enhance workplace safety. The most successful example of a settings approach is the US Air Force suicide prevention initiative (Knox et al. 2003). The workplace is an ideal setting for health promotion and suicide prevention initiatives: it provides regular and easy access to a large number of people who make up a relatively stable population and it may encourage sustained peer support and positive peer pressure (Tones & Tilford, 1994). Stakeholders of programs and evaluations should recognize that workplaces and communities are dynamic, and socially, culturally, and economically heterogeneous; therefore, strategies need to be adapted to local needs and possibilities. Effective workplace suicide prevention initiatives depend, in part, on the interest and willingness of employers to support such programs and on employees to participate. Finding ways to make suicide prevention programs relevant and acceptable to employees is an important factor to their
effectiveness. Workplace health promotion guidelines suggest that programs need to feed into different levels to facilitate sustained behavioural change. At the individual level, a program should incorporate a range of educational strategies; at the organizational level, support mechanisms throughout the organization should reinforce and encourage positive actions (e.g. intervening with distressed individuals). Finally, at the community level, workplace suicide prevention initiatives should have the potential to be actively disseminated by employees to their families and social networks (Wilson, Holman & Hammock, 1996). However, research is limited on the long-term effects on individuals expose suicide prevention training (Isaac et al., 2009). There is an urgent need for more rigorous studies to determine the immediate and longitudinal effects of gatekeepers’ training on knowledge acquisition, intervention skills and referrals to appropriate community resources. Moreover, rigorous research is needed to demonstrate whether gatekeeper training programs increase help-seeking behaviours of at-risk individuals and ultimately have an impact on decreasing the rates of suicide ideation, attempts and deaths by suicide.

Based on this program, evaluation the following recommendations are suggested as strategies for maintaining TTC employees’ gatekeeper knowledge, intervention skills and motivation to intervene with distressed patrons. Recommendations for future gatekeeper program development and suicide prevention research activities are also provided.

**Recommendation 1:**

Gatekeeper refresher course for employees to periodically attend a short version of suicide prevention training that highlights the basics of suicide prevention knowledge and skills. The gatekeeper refresher can be offered in different formats such as face-to-face or on-line. The delivery model of the refresher course needs to be adapted to meet diverse learning styles, support the needs of participants, and to effectively manage the rhythm and pace of delivery of such an intense learning experience.

**Recommendation 2:**

A periodic newsletter via email or regular mail would include updates of suicide prevention at the TTC, reminders of refresher workshops, new national and regional suicide statistics, new
procedures, tips, success stories, and/or a question and answer portion. The newsletter can incorporate other mental health promotion activities such as raising awareness about depression or substance abuse.

**Recommendation 3:**
To be effective and efficient, implementation of training programs must start with a training needs assessment. Performing training needs assessment is a critical activity for the training and development of appropriate programs. The results of the needs assessment would allow to set specific training objectives and provide clear guidelines for the development of gatekeeper programs which meet particular occupational group’s needs (e.g. Special Constables) and are appropriate and congruent with their job related tasks and activities.

**Recommendation 4:**
Further research should be conducted into the barriers inhibiting employees from intervening with distressed at-risk for suicide patrons. This information could be utilised to refine or design future suicide intervention training programs and evaluative research.

**Recommendation 5:**
Given the low rate of suicide in the Toronto subway system, there is a need to collect surrogate endpoint data to substitute for suicides. Surrogate endpoints frequently substitute for rare outcomes in research. The surrogate endpoints will allow the TTC to measure the effect of the gatekeeper program on reducing subway suicides by investigating more readily available proximate outcomes such as Mental Health Act (MHA) apprehensions of distressed patrons. Treating MHA apprehensions as surrogate endpoints would serve as a proxy for suicides that occur in the subway system and potentially provide insight into the phenomenon of subway suicide. Thus, it is critical that TTC management develop appropriate standardized incident report forms that allow the collection of pertinent surrogate data about suicidality in the subway system.

**Concluding Remarks**
This study evaluated the suicide prevention training implemented at the Toronto Transit Commission using a comprehensive analytical framework which tested various aspects of conducting such an evaluation. Results of this study illustrated that safeTALK and suicideAWARE implemented at the TTC for the commission’s employees had immediate and long-term effects on program participants. The training increased participants’ knowledge of suicide and suicidal behaviour, enhanced positive attitudes toward the suicidal individual, suicide intervention, and improved intervention skills. These findings support all hypothesized outcomes. The study also examined and appraised the Kirkpatrick Model as an analytical framework for suicide prevention program evaluation. The empirical findings from this study support the premise that the Model could be adapted for use in gatekeeper program evaluations. The Model provides a highly relevant, well-rounded, rigorous approach to suicide prevention program evaluations.
REFERENCES


Galea, S., Tracy, M., Piper, T. M., Bucciarelli, A. M., Tardiff, K., Gershon, R. et al. (2006, November). Epidemiology of suicide in the New York City subway system. Recorded presentation at APHA 134th Annual Meeting and Exposition, Boston, MA.


**Web Sources**

http://www.livingworks.net/LW.php3/24/06.
http://www.minddisorders.com/Flu=Inv/Intervention.html 4/5/06.
http://www.e-law.gov.on.ca/DBlaws/Statutes/English/90m07_e.htm 4/5/06.
APPENDIX A:
The Collaborative Project and Development of a Comprehensive List of Warning Signs
Phase 2: The Collaborative Project

Collaboration between the Toronto Transit Commission (TTC) and the Arthur Sommer Rotenberg Chair in Suicide Studies - St. Michael’s Hospital and the University of Toronto - on the issue of suicide prevention has been ongoing for almost the last decade. In June of 2004, a collaboration agreement was signed between the Arthur Sommer Rotenberg Chair in Suicide Studies, at St. Michael’s Hospital, and the Toronto Transit Commission (TTC) for the purpose of assisting the TTC in the development of a new gatekeeper educational suicide prevention program aimed at reducing the incidences of suicide attempts and completed suicides on the Toronto subway system. The Overview Logic Model (Appendix B) illustrates the resources, activities, participation populations and expected outcomes (short, medium, and long-term) for the Collaborative Project.

It was rationalized that alertness to warning signs, symptoms and suicidal behaviour can be fostered by training (Living Works, 1999) and effective referral and intervention strategies can be implemented by anyone who has an interest in helping suicidal individuals (Scott & Armson, 2000).

The over-arching objective of the collaborative project was to prevent suicide attempts and completed suicides in the Toronto subway system.

The specific objectives of the Collaboration Project were:

1. To assist the TTC in improving their existing suicide prevention program.
2. To develop operational criteria for the identification of distressed and at-risk persons using the TTC. The criteria are to be incorporated into the TTC gatekeeper educational suicide prevention program.

1. The Arthur Sommer Rotenberg Chair in Suicide Studies and the TTC will participate in a site visit to Montreal to meet with the Montreal Metro and Dr. Brian Mishara to learn about their progress and to examine opportunities for collaboration.
2. The Arthur Sommer Rotenberg Chair in Suicide Studies will assist in the development of criteria to identify distressed and at-risk persons using the TTC.
3. The Arthur Sommer Rotenberg Chair in Suicide Studies will assist in the development of an effective intervention protocol which reduces risks to the passengers, does not
4. increase the risk to TTC staff and permits a graceful withdrawal in case of a false positive. The Arthur Sommer Rotenberg Chair in Suicide Studies will assist in the development of a standard report of suicide incidents.

5. The Arthur Sommer Rotenberg Chair in Suicide Studies will act in an advisory role in the development and implementation of a training program for staff.

6. The Arthur Sommer Rotenberg Chair in Suicide Studies will help identify a qualified training agency to deliver the training.

7. The Arthur Sommer Rotenberg Chair in Suicide Studies will develop an approach to an evaluation project to determine the success of the employee awareness program, both from the employees’ perspective, and from the people who have received the interventions resultant from the program.

8. The Arthur Sommer Rotenberg Chair in Suicide Studies will assist in determining possible requirements and sources for external funding for the future development of the evaluation project as required.

The initial step in implementation was the formation of a steering committee representing the TTC and researchers from the Arthur Sommer Rotenberg Chair. The committee began its work in the summer of 2004. All of the outlined collaborative project goals were achieved.
APPENDIX B:
The Development of a Comprehensive List of Warning Signs
Phase 3: Development of a Comprehensive Warning Signs List

Implicit in the general suicide education approach is recognition of the difficulty of determining who, among millions of healthy patrons, is truly at high risk of suicide. The identification of the specific populations at risk for subway suicide has several significant and valuable implications for devising strategies for prevention. The detection of the particular characteristics of the populations at risk is useful in the enrichment of a gatekeeper training program for employees of the Toronto Transit Commission (TTC). Identification of suicidal individuals, a primary component of gatekeeper training, may be strengthened by the astute observations of the TTC employees themselves. The survey conducted specifically targeted key informants from among TTC employees to solicit their expert opinions on warning signs exhibited by individuals who have attempted or completed suicide in the Toronto subway system.

To facilitate the identification of patrons at risk for suicide, we sought to establish a comprehensive list of behavioural cues exhibited by distressed and at-risk-of-suicide patrons. The completeness and clarity of the list was validated with three stakeholder groups: mental health personnel working with the St. Michael’s Hospital Psychiatric Emergency Service, TTC Special Constables, Mobile and Surface Supervisors, and Train Operators.

To develop this comprehensive list of 20 behavioural cues the following strategies were used:

- An extensive literature review was conducted and published in the Journal of Urban Health (Ratnayake, Links, &. Eynan, 2007).
- The Montreal Metro pamphlet was translated from French to English and symptoms agitated and distressed individuals may exhibit were added to the list.
- The expert opinions of clinicians experienced in Emergency Psychiatry Service at St. Michael’s Hospital were sought.
- The Clinicians’ comments were incorporated in the survey questionnaire designed for TTC Special Constables and Mobile and Surface Operators and the interview guide used with Train Operators.
TTC Special Constables, Mobile and Surface Supervisors were surveyed, and Train Operators who had experienced a suicide of a patron were interviewed.

The list of behavioural cues, comments, and observations made by TTC Special Constables, Mobile and Surface Supervisors, Train Operators and Train Guards was shared with LivingWorks who incorporated the behaviours into vignettes developed specifically for the TTC and integrated in the training program developed.
APPENDIX C:
The Logic Models for the TTC Program Evaluation
<table>
<thead>
<tr>
<th>Input</th>
<th>Activities</th>
<th>Output</th>
<th>Immediate Outcomes</th>
<th>Intermediate Outcomes</th>
<th>Ultimate Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Funds allocated by the TTC for the implementation of a suicide prevention program</td>
<td>Gatekeeper Program Implementation for TTC employees</td>
<td>Deliver safeTALK suicideAWARE training to select groups of employees</td>
<td>Knowledge Skills and Positive attitude are enhanced</td>
<td>Participants used their knowledge and skills in workplace context</td>
<td>Participants are more effective in identifying distressed patrons Lower suicide rates</td>
</tr>
</tbody>
</table>
Rational for Training Programs
Process and Outcome

Gatekeeper Training
for
TTC Special Constables
Supervisors
Train Operators

↓
Attitudes about suicide
Attitudes about suicidal person
Knowledge of suicidal behaviour
Development & enhancement of intervention skills
Knowledge of community resources

↓
Increase awareness of attitude barriers
Increase factual knowledge about suicide
Facilitate recognition of at-risk individuals
Develop Risk Assessment skills
Increase confidence & competence to intervene
Motivate to network with other resources

↓
Increase desirable attitudes
Increase factual knowledge about suicide
Increase likelihood staff will report distress individual
Increase confidence and competence to intervene
Increase likelihood employees will intervene
Access and initiate appropriate referrals

↓
Early preventative interventions
Engagement of social supports
Referral to treatment

↓
Reduce incidents of suicidal behaviour & injuries in the subway system

↓
Fewer suicides
APPENDIX D:

suicideAWARE pre-training Questionnaire
Completion of this questionnaire is voluntary. If you do not wish to complete the questionnaire you do not have to answer the questions.

1. My TTC employee number is: ________

2. My gender (check one):
   □ Male    □ Female

3. My Age (check one):
   □ 18-24 years old
   □ 25-34 years old
   □ 35-49 years old
   □ 50-65 years old

4. Position at TTC (check one):
   □ Special Constable    □ Subway Supervisor
   □ Mobile Supervisor    □ Chief Supervisor
   □ Surface Supervisor   □ Other: (specify) ________________________
   □ Train Operator

5. Years of employment with the TTC (check one):
   □ Less than one year
   □ 1-2 years
   □ 3-5 years
   □ 6-10 years
   □ 10 + years

6. Number of contacts you had with a suicidal person:
   a) In all your years at the TTC:_____
   b) In the past year:_____
   c) In the past month:_____
This is not a test but a survey of your opinions; there is no right or wrong answer, only your honest opinion counts. Please check one box:

1. Almost everyone at one time thought about suicide.
   - [ ] Strongly agree  [ ] Agree  [ ] Undecided  [ ] Disagree  [ ] Strongly disagree

2. Most suicidal attempts are impulsive in nature.
   - [ ] Strongly agree  [ ] Agree  [ ] Undecided  [ ] Disagree  [ ] Strongly disagree

3. Those who threaten suicide rarely do so.
   - [ ] Strongly agree  [ ] Agree  [ ] Undecided  [ ] Disagree  [ ] Strongly disagree

4. Suicide is more prevalent among the very rich and the very poor.
   - [ ] Strongly agree  [ ] Agree  [ ] Undecided  [ ] Disagree  [ ] Strongly disagree

5. Most people who try to kill themselves don’t really want to die.
   - [ ] Strongly agree  [ ] Agree  [ ] Undecided  [ ] Disagree  [ ] Strongly disagree

6. Suicide happens without warning.
   - [ ] Strongly agree  [ ] Agree  [ ] Undecided  [ ] Disagree  [ ] Strongly disagree

7. About 80% of those who commit suicide have attempted suicide at least once before.
   - [ ] Strongly agree  [ ] Agree  [ ] Undecided  [ ] Disagree  [ ] Strongly disagree

8. It’s rare for someone who is thinking about suicide to be dissuaded by a ‘friendly ear.’
   - [ ] Strongly agree  [ ] Agree  [ ] Undecided  [ ] Disagree  [ ] Strongly disagree

9. A large percentage of suicide victims come from a broken home.
   - [ ] Strongly agree  [ ] Agree  [ ] Undecided  [ ] Disagree  [ ] Strongly disagree
10. The possibility of committing suicide is greater for older people (60 and over) than for young people (20 to 30 years).

□ Strongly agree □ Agree □ Undecided □ Disagree □ Strongly disagree

11. Once a person is suicidal, s/he is suicidal forever.

□ Strongly agree □ Agree □ Undecided □ Disagree □ Strongly disagree

12. Improvement following a suicidal crisis indicates that the risk is over.

□ Strongly agree □ Agree □ Undecided □ Disagree □ Strongly disagree

13. Once a person pursues a suicidal attempt, the probability of her/his trying again is minimal.

□ Strongly agree □ Agree □ Undecided □ Disagree □ Strongly disagree

14. Suicide attempters who use public places (such as a subway, a bridge) are more interested in getting attention.

□ Strongly agree □ Agree □ Undecided □ Disagree □ Strongly disagree

15. If someone wants to commit suicide, it is their business, and we should not interfere.

□ Strongly agree □ Agree □ Undecided □ Disagree □ Strongly disagree

16. A suicide attempt is essentially a “cry for help.”

□ Strongly agree □ Agree □ Undecided □ Disagree □ Strongly disagree

17. Usually, relatives of a suicide victim had no idea of what was about to happen.

□ Strongly agree □ Agree □ Undecided □ Disagree □ Strongly disagree

18. Those people who attempt suicide are usually trying to get sympathy from others.

□ Strongly agree □ Agree □ Undecided □ Disagree □ Strongly disagree

19. Potentially, every one of us can be a suicide victim.

□ Strongly agree □ Agree □ Undecided □ Disagree □ Strongly disagree
20. Most suicide victims are older persons with little to live for.

☐ Strongly agree  ☐ Agree  ☐ Undecided  ☐ Disagree  ☐ Strongly disagree
TTC SUICIDE PREVENTION TRAINING PROGRAM

SUICIDE-RISK PROCEDURAL QUESTIONNAIRE

1. Which of the following is a good indicator that the passenger is at risk:
   a. They let several trains go by and they look depressed or distressed
   b. They are at the far end of the platform (train entrance point)
   c. They are agitated and pace at the far end of the platform close to the entrance to the tunnel
   d. They stand on the yellow line and they look depressed or distressed
   e. All of the above

2. Transit control should be called if a passenger on the platform:
   a. Extends part of their body beyond the platform
   b. Argues or fights with strangers
   c. Sits on the edge of the platform and dangle their feet over the rail
   d. Skateboards, rollerblades
   e. All of the above

3. Which of the following is a warning sign and help should be called:
   a. Passenger sits on the bench and cries
   b. Passenger takes off their clothes or removes their shoes
   c. Passenger says they are going to end their lives
   d. Passenger stands close to a foot bridge and appears depressed
   e. All of the above

4. Transit control should be called if a passenger:
   a. is inappropriately dressed for the weather
   b. is wearing army fatigues
   c. is at track level
   d. carries an umbrella on a sunny day
   e. none of the above
APPENDIX E:
safeTALK pre-training Questionnaire
Completion of this questionnaire is voluntary. If you do not wish to complete the questionnaire you do not have to answer the questions.

1. My TTC employee number is: ________

2. My gender (check one):
   - □ Male
   - □ Female

3. My Age (check one):
   - □ 18-24 years old
   - □ 25-34 years old
   - □ 35-49 years old
   - □ 50-65 years old

4. Position at TTC (check one):
   - □ Special Constable
   - □ Mobile Supervisor
   - □ Surface Supervisor
   - □ Subway Supervisor
   - □ Chief Supervisor
   - □ Other: (specify) __________________________
   - □ Train Operator

5. Years of employment with the TTC (check one):
   - □ Less than one year
   - □ 1-2 years
   - □ 3-5 years
   - □ 6-10 years
   - □ 10 + years

We would like to know your level of experience with Mental Health Act (MHA) apprehensions and suicidal patrons.

6. Number of times you were involved in a MHA apprehension:
   - a) In all your years of employment at the TTC:_____
   - b) In the last year:_____
   - c) In the last month:_____
   - d) Does not apply: _____

7. Number of contacts you had with a suicidal person:
   - a) In all your years at the TTC:_____
   - b) In the past year:_____
   - c) In the past month:_____
**TTC SUICIDE PREVENTION TRAINING PROGRAM**

**Suicide Opinion Questionnaire**

This is **not** a test but a survey of your opinions; there is no right or wrong answer, only your honest opinion counts. Please check one box:

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
</table>
| 1. Almost everyone at one time thought about suicide.  
   □ Strongly agree  □ Agree  □ Undecided  □ Disagree  □ Strongly disagree |
| 2. Most suicidal attempts are impulsive in nature.  
   □ Strongly agree  □ Agree  □ Undecided  □ Disagree  □ Strongly disagree |
| 3. Those who threaten suicide rarely do so.  
   □ Strongly agree  □ Agree  □ Undecided  □ Disagree  □ Strongly disagree |
| 4. Suicide is more prevalent among the very rich **and** the very poor.  
   □ Strongly agree  □ Agree  □ Undecided  □ Disagree  □ Strongly disagree |
| 5. Most people who try to kill themselves don’t really want to die.  
   □ Strongly agree  □ Agree  □ Undecided  □ Disagree  □ Strongly disagree |
| 6. Suicide happens without warning.  
   □ Strongly agree  □ Agree  □ Undecided  □ Disagree  □ Strongly disagree |
| 7. About 80% of those who commit suicide have attempted suicide at least once before.  
   □ Strongly agree  □ Agree  □ Undecided  □ Disagree  □ Strongly disagree |
| 8. It’s rare for someone who is thinking about suicide to be dissuaded by a ‘friendly ear.’  
   □ Strongly agree  □ Agree  □ Undecided  □ Disagree  □ Strongly disagree |
| 9. A large percentage of suicide victims come from a broken home.  
   □ Strongly agree  □ Agree  □ Undecided  □ Disagree  □ Strongly disagree |
10. The possibility of committing suicide is greater for older people (60 and over) than for young people (20 to 30 years).

☐ Strongly agree  ☐ Agree  ☐ Undecided  ☐ Disagree  ☐ Strongly disagree

11. Once a person is suicidal, s/he is suicidal forever.

☐ Strongly agree  ☐ Agree  ☐ Undecided  ☐ Disagree  ☐ Strongly disagree

12. Improvement following a suicidal crisis indicates that the risk is over.

☐ Strongly agree  ☐ Agree  ☐ Undecided  ☐ Disagree  ☐ Strongly disagree

13. Once a person pursues a suicidal attempt, the probability of her/his trying again is minimal.

☐ Strongly agree  ☐ Agree  ☐ Undecided  ☐ Disagree  ☐ Strongly disagree

14. Suicide attempters who use public places (such as a subway, a bridge) are more interested in getting attention.

☐ Strongly agree  ☐ Agree  ☐ Undecided  ☐ Disagree  ☐ Strongly disagree

15. If someone wants to commit suicide, it is their business, and we should not interfere.

☐ Strongly agree  ☐ Agree  ☐ Undecided  ☐ Disagree  ☐ Strongly disagree

16. A suicide attempt is essentially a “cry for help.”

☐ Strongly agree  ☐ Agree  ☐ Undecided  ☐ Disagree  ☐ Strongly disagree

17. Usually, relatives of a suicide victim had no idea of what was about to happen.

☐ Strongly agree  ☐ Agree  ☐ Undecided  ☐ Disagree  ☐ Strongly disagree

18. Those people who attempt suicide are usually trying to get sympathy from others.

☐ Strongly agree  ☐ Agree  ☐ Undecided  ☐ Disagree  ☐ Strongly disagree

19. Potentially, every one of us can be a suicide victim.

☐ Strongly agree  ☐ Agree  ☐ Undecided  ☐ Disagree  ☐ Strongly disagree
20. Most suicide victims are older persons with little to live for.

☐ Strongly agree  ☐ Agree  ☐ Undecided  ☐ Disagree  ☐ Strongly disagree
TTC SUICIDE PREVENTION TRAINING PROGRAM

SUICIDE-RISK PROCEDURAL QUESTIONNAIRE

1. Which of the following is a good indicator that the passenger is at risk:
   
   a. They let several trains go by and they look depressed or distressed
   b. They are at the far end of the platform (train entrance point)
   c. They are agitated and pace at the far end of the platform close to the entrance to the tunnel
   d. They stand on the yellow line and they look depressed or distressed
   e. All of the above

2. Transit control should be called if a passenger on the platform:
   
   a. Extends part of their body beyond the platform
   b. Argues or fights with strangers
   c. Sits on the edge of the platform and dangle their feet over the rail
   d. Skateboards, rollerblades
   e. All of the above

3. Which of the following is a warning sign and help should be called:
   
   a. Passenger sits on the bench and cries
   b. Passenger takes off their clothes or removes their shoes
   c. Passenger says they are going to end their lives
   d. Passenger stands close to a foot bridge and appears depressed
   e. All of the above

4. Transit control should be called if a passenger:
   
   a. is inappropriately dressed for the weather
   b. is wearing army fatigues
   c. is at track level
   d. carries an umbrella on a sunny day
   e. none of the above
Part I Multiple Choice: For the questions which follow, select the best response by circling the appropriate letter. If uncertain, feel free to guess.

1. Suicide is most likely a result of:
   a. overwhelming stress
   b. clinical depression
   c. substance abuse
   d. no single cause

2. When confronted with the possibility of suicidal behaviour in a person, a helper should immediately:
   a. refer the person to experienced suicide resources
   b. discuss it directly with the person
   c. call in significant others in the person's life
   d. encourage the person to talk about positive aspects of his or her life

3. Suicide plans are assessed on the basis of a person's:
   a. level of planning
   b. age
   c. stated seriousness
   d. apparent distress

4. People who express suicide intentions:
   a. clearly want to die
   b. are ambivalent about dying
   c. want to punish others
   d. are manipulating

5. Of the following, which provides the most important information in assessing the risk of suicide:
   a. symptoms
   b. stress
   c. resources
   d. physical health

6. What generally determines if behaviour is suicidal?
   a. the mood of the person
   b. the intent of the person
   c. the lethality of the method used
   d. the history of prior behaviour

7. Of the following, which provides the least important information in assessing the risk of suicide?
   a. symptoms
   b. stress
   c. resources
   d. physical health
8. Stress in people lives is:
   a. most commonly a subjective matter
   b. determined by external events
   c. synonymous with crisis
   d. highly related to level of suicide risk

9. If someone is showing signs and symptoms suggestive of suicide, a helper should:
   a. gather more information about what is bothering the person
   b. inquire about the support available from family and friends
   c. determine if substance abuse is a factor
   d. ask if the person is thinking about suicide

10. If someone admits to feeling suicidal, a helper should:
   a. calmly inquire about what is happening in their life
   b. find out if they’ve thought of how they would do it
   c. inform significant others
   d. arrange immediate referral

11. Which of the following phase comprise the safeTALK Model used in the Suicide Intervention Training Program:
   a. Invitation; tell, ask, listen, keeping safe
   b. Tell, invitation, listen, ask, keeping safe
   c. Ask, listen, invitations, tell, keep safe
   d. Keep safe, invitation, tell, ask, listen

12. Which of the following is considered an invitation for help:
   a. Actions
   b. Feelings
   c. Thoughts
   d. Reactions
   e. Life situations
   f. All of the above
   g. None of the above

13. Key tasks in the safeTALK Suicide Intervention Model are:
   a. ask, listen, keeping safe
   b. keeping safe, ask, listen
   c. listen, keeping safe, ask
   d. none of the above

14. It is important to move through a suicide intervention process while (circle all that apply):
   a. keeping self and others safe by securing the platform
   b. ensuring trains return to service quickly
   c. communicating with transit control
   d. all of the above
   e. none of the above

15. What will be your role in suicide intervention:
   a. engage in counseling
   b. refer to appropriate resources
   c. call significant others
   d. insure the person gets on the next train
16. Helper attitude to someone at risk of suicide:
   a. should be made clear to a person at risk of suicide
   b. should be kept from person at risk of suicide
   c. may have either positive or negative effects
   d. have little to do with the process of intervention

17. Which of the following action plans would likely be most suitable for someone at risk of suicide:
   a. Agreement for therapy
   b. agreement to meet again
   c. agreement to call a crisis line if troubled again with thoughts of suicide
   d. agreement to talk with significant other
   e. referral to appropriate resources

Part II

The following items represent a series of excerpts from intervention situations. Each excerpt begins with an expression by the patron concerning some aspect of the situation s/he faces, followed by two possible helper responses to the patron’s remark. You are to rate the response in term of how appropriate or inappropriate you feel the reply is to the patron’s comment. In the blank you should record a rating from -3 to +3, corresponding to the chart below. Be sure to respond to each item, and try not to leave any blanks.

+3 = Highly appropriate response
+2 = Appropriate response
+1 = Marginally appropriate response
0 = Neither appropriate nor inappropriate
-1 = Marginally inappropriate response
-2 = Inappropriate response
-3 = Highly inappropriate response

1. Patron 1:
   I decided to come down to the subway station tonight because I really feel things are too much for me……. Now my health is going downhill too, on top of all the rest. Without my husband around to care for me anymore, it just seems like the end of the world.

   _______ Helper A: Try not to worry too much about things. Everything will be all right.
   _______ Helper B: You must feel pretty lonely and afraid of what might happened… Are you thinking about suicide?

2. Patron 2:
   My life has been worthless since my wife, Emma, died four years ago. The kids are grown and married now, and I’ve been retired from my job at the railroad for some times. It just seems that I’d be better off dead.

   _______ Helper A: But try to think of what Emma would want for you. She would want you to continue leading a productive life, wouldn’t she?
Helper B: It sounds like everything just collapsed around you since Emma died ... But what happened recently to make things even worse, to make you think dying is the only way out?

3. Patron 3:
How could you help me? Have you ever wanted to kill yourself?

Helper A: You sound like you are concern about whether I can understand and help you.

Helper B: Sure I have thought about suicide sometimes. But I also found more realistic solutions to my problems.

4. Patron 4:
I am so lonely so tired (crying). There just isn’t any where left to turn.

Helper A: You seem so alone, so miserable. Have you been feeling suicidal?

Helper B: Come on now. Things can’t be all that bad.

5. Patron 5:
Why should you care about me anyway?

Helper A: I’ve been trained to care about people. That’s my job.

Helper B: Because I think your death will be a terrible waste, and it concerns me that things are so that you are considering suicide. You need help to get through this critical period.

6. Patron 6:
No one can understand the kind of pain I’ve been going through. Sometimes I feel like I have to hurt myself.....

Helper A: It seems that you’ve been suffering so much that hurting yourself is the only you can make the pain go away.

Helper B: But you are so young, you have so much to live for. How can you think of killing yourself?
APPENDIX F:

suicideAWARE post-training Questionnaire
 Completion of this questionnaire is voluntary. If you do not wish to complete the questionnaire you do not have to answer the questions. This is not a test but a survey of your opinions; there is no right or wrong answer, only your honest opinion counts. Please check one box:

1. Almost everyone at one time thought about suicide.
   - [ ] Strongly agree  [ ] Agree  [ ] Undecided  [ ] Disagree  [ ] Strongly disagree

2. Most suicidal attempts are impulsive in nature.
   - [ ] Strongly agree  [ ] Agree  [ ] Undecided  [ ] Disagree  [ ] Strongly disagree

3. Those who threaten suicide rarely do so.
   - [ ] Strongly agree  [ ] Agree  [ ] Undecided  [ ] Disagree  [ ] Strongly disagree

4. Suicide is more prevalent among the very rich and the very poor.
   - [ ] Strongly agree  [ ] Agree  [ ] Undecided  [ ] Disagree  [ ] Strongly disagree

5. Most people who try to kill themselves don't really want to die.
   - [ ] Strongly agree  [ ] Agree  [ ] Undecided  [ ] Disagree  [ ] Strongly disagree

6. Suicide happens without warning.
   - [ ] Strongly agree  [ ] Agree  [ ] Undecided  [ ] Disagree  [ ] Strongly disagree

7. About 80% of those who commit suicide have attempted suicide at least once before.
   - [ ] Strongly agree  [ ] Agree  [ ] Undecided  [ ] Disagree  [ ] Strongly disagree

8. It’s rare for someone who is thinking about suicide to be dissuaded by a ‘friendly ear.’
   - [ ] Strongly agree  [ ] Agree  [ ] Undecided  [ ] Disagree  [ ] Strongly disagree

9. A large percentage of suicide victims come from a broken home.
   - [ ] Strongly agree  [ ] Agree  [ ] Undecided  [ ] Disagree  [ ] Strongly disagree

10. The possibility of committing suicide is greater for older people (60 and over) than for young people (20 to 30 years).
11. Once a person is suicidal, s/he is suicidal forever.

12. Improvement following a suicidal crisis indicates that the risk is over.

13. Once a person pursues a suicidal attempt, the probability of her/his trying again is minimal.

14. Suicide attempters who use public places (such as a subway, a bridge) are more interested in getting attention.

15. If someone wants to commit suicide, it is their business, and we should not interfere.

16. A suicide attempt is essentially a “cry for help.”

17. Usually, relatives of a suicide victim had no idea of what was about to happen.

18. Those people who attempt suicide are usually trying to get sympathy from others.

19. Potentially, every one of us can be a suicide victim.

20. Most suicide victims are older persons with little to live for.
TTC SUICIDE PREVENTION TRAINING PROGRAM

SUICIDE-RISK PROCEDURAL QUESTIONNAIRE

1. Which of the following is a good indicator that the passenger is at risk:
   a. They let several trains go by and they look depressed or distressed
   b. They are at the far end of the platform (train entrance point)
   c. They are agitated and pace at the far end of the platform close to the entrance to the tunnel
   d. They stand on the yellow line and they look depressed or distressed
   e. All of the above

2. Transit control should be called if a passenger on the platform:
   a. Extends part of their body beyond the platform
   b. Argues or fights with strangers
   c. Sits on the edge of the platform and dangle their feet over the rail
   d. Skateboards, rollerblades
   e. All of the above

3. Which of the following is a warning sign and help should be called:
   a. Passenger sits on the bench and cries
   b. Passenger takes off their clothes or removes their shoes
   c. Passenger says they are going to end their lives
   d. Passenger stands close to a foot bridge and appears depressed
   e. All of the above

4. Transit control should be called if a passenger:
   a. is inappropriately dressed for the weather
   b. is wearing army fatigues
   c. is at track level
   d. carries an umbrella on a sunny day
   e. none of the above
APPENDIX G:
safeTALK post-training Questionnaire
safeTALK
TTC SUICIDE PREVENTION TRAINING PROGRAM
Suicide Opinion Questionnaire
(Post-training)

My TTC Employee Number is: ___________________

Completion of this questionnaire is voluntary. If you do not wish to complete the questionnaire you do not have to answer the questions.
This is not a test but a survey of your opinions; there is no right or wrong answer, only your honest opinion counts. Please check one box:

1. Almost everyone at one time thought about suicide.
   □ Strongly agree  □ Agree  □ Undecided  □ Disagree  □ Strongly disagree

2. Most suicidal attempts are impulsive in nature.
   □ Strongly agree  □ Agree  □ Undecided  □ Disagree  □ Strongly disagree

3. Those who threaten suicide rarely do so.
   □ Strongly agree  □ Agree  □ Undecided  □ Disagree  □ Strongly disagree

4. Suicide is more prevalent among the very rich and the very poor.
   □ Strongly agree  □ Agree  □ Undecided  □ Disagree  □ Strongly disagree

5. Most people who try to kill themselves don’t really want to die.
   □ Strongly agree  □ Agree  □ Undecided  □ Disagree  □ Strongly disagree

6. Suicide happens without warning.
   □ Strongly agree  □ Agree  □ Undecided  □ Disagree  □ Strongly disagree

7. About 80% of those who commit suicide have attempted suicide at least once before.
   □ Strongly agree  □ Agree  □ Undecided  □ Disagree  □ Strongly disagree

8. It’s rare for someone who is thinking about suicide to be dissuaded by a ‘friendly ear.’
   □ Strongly agree  □ Agree  □ Undecided  □ Disagree  □ Strongly disagree

9. A large percentage of suicide victims come from a broken home.
   □ Strongly agree  □ Agree  □ Undecided  □ Disagree  □ Strongly disagree
10. The possibility of committing suicide is greater for older people (60 and over) than for young people (20 to 30 years).
   □ Strongly agree □ Agree □ Undecided □ Disagree □ Strongly disagree

11. Once a person is suicidal, s/he is suicidal forever.
   □ Strongly agree □ Agree □ Undecided □ Disagree □ Strongly disagree

12. Improvement following a suicidal crisis indicates that the risk is over.
   □ Strongly agree □ Agree □ Undecided □ Disagree □ Strongly disagree

13. Once a person pursues a suicidal attempt, the probability of her/his trying again is minimal.
   □ Strongly agree □ Agree □ Undecided □ Disagree □ Strongly disagree

14. Suicide attempters who use public places (such as a subway, a bridge) are more interested in getting attention.
   □ Strongly agree □ Agree □ Undecided □ Disagree □ Strongly disagree

15. If someone wants to commit suicide, it is their business, and we should not interfere.
   □ Strongly agree □ Agree □ Undecided □ Disagree □ Strongly disagree

16. A suicide attempt is essentially a “cry for help.”
   □ Strongly agree □ Agree □ Undecided □ Disagree □ Strongly disagree

17. Usually, relatives of a suicide victim had no idea of what was about to happen.
   □ Strongly agree □ Agree □ Undecided □ Disagree □ Strongly disagree

18. Those people who attempt suicide are usually trying to get sympathy from others.
   □ Strongly agree □ Agree □ Undecided □ Disagree □ Strongly disagree

19. Potentially, every one of us can be a suicide victim.
   □ Strongly agree □ Agree □ Undecided □ Disagree □ Strongly disagree

20. Most suicide victims are older persons with little to live for.
   □ Strongly agree □ Agree □ Undecided □ Disagree □ Strongly disagree
TTC SUICIDE PREVENTION TRAINING PROGRAM

SUICIDE-RISK PROCEDURAL QUESTIONNAIRE

1. Which of the following is a good indicator that the passenger is at risk:
   a. They let several trains go by and they look depressed or distressed
   b. They are at the far end of the platform (train entrance point)
   c. They are agitated and pace at the far end of the platform close to the entrance to the tunnel
   d. They stand on the yellow line and they look depressed or distressed
   e. All of the above

2. Transit control should be called if a passenger on the platform:
   a. Extends part of their body beyond the platform
   b. Argues or fights with strangers
   c. Sits on the edge of the platform and dangle their feet over the rail
   d. Skateboards, rollerblades
   e. All of the above

3. Which of the following is a warning sign and help should be called:
   a. Passenger sits on the bench and cries
   b. Passenger takes off their clothes or removes their shoes
   c. Passenger says they are going to end their lives
   d. Passenger stands close to a foot bridge and appears depressed
   e. All of the above

4. Transit control should be called if a passenger:
   a. is inappropriately dressed for the weather
   b. is wearing army fatigues
   c. is at track level
   d. carries an umbrella on a sunny day
   e. none of the above
TTC SUICIDE PREVENTION TRAINING
EVALUATION

Intervention Knowledge Test

Part I Multiple Choice: For the questions which follow, select the best response by circling the appropriate letter. If uncertain, feel free to guess.

1. Suicide is most likely a result of:
   a. overwhelming stress
   b. clinical depression
   c. substance abuse
   d. no single cause

2. When confronted with the possibility of suicidal behaviour in a person, a helper should immediately:
   a. refer the person to experienced suicide resources
   b. discuss it directly with the person
   c. call in significant others in the persons life
   d. encourage the person to talk about positive aspects of his or her life

3. Suicide plans are assessed on the basis of a persons:
   a. level of planning
   b. age
   c. stated seriousness
   d. apparent distress

4. People who express suicide intentions:
   a. clearly want to die
   b. are ambivalent about dying
   c. want to punish others
   d. are manipulating

5. Of the following, which provides the most important information in assessing the risk of suicide:
   a. symptoms
   b. stress
   c. resources
   d. physical health

6. What generally determines if behaviour is suicidal?
   a. the mood of the person
   b. the intent of the person
   c. the lethality of the method used
   d. the history of prior behaviour

7. Of the following, which provides the least important information in assessing the risk of suicide?
   a. symptoms
   b. stress
   c. resources
   d. physical health
9. **Stress in people lives is:**
   a. most commonly a subjective matter
   b. determined by external events
   c. synonymous with crisis
   d. highly related to level of suicide risk

10. **If someone is showing signs and symptoms suggestive of suicide, a helper should:**
    a. gather more information about what is bothering the person
    b. inquire about the support available from family and friends
    c. determine if substance abuse is a factor
    d. ask if the person is thinking about suicide

11. **If someone admits to feeling suicidal, a helper should:**
    a. calmly inquire about what is happening in their life
    b. find out if they’ve thought of how they would do it
    c. inform significant others
    d. arrange immediate referral

12. **Which of the following phase comprise the safeTALK Model used in the Suicide Intervention Training Program:**
    a. Invitation; tell, ask, listen, keeping safe
    b. Tell, invitation, listen, ask, keeping safe
    c. Ask, listen, invitations, tell, keep safe
    d. Keep safe, invitation, tell, ask, listen

13. **Which of the following is considered an invitation for help:**
    a. Actions
    b. Feelings
    c. Thoughts
    d. Reactions
    e. Life situations
    f. All of the above
    g. None of the above

14. **Key tasks in the safeTALK Suicide Intervention Model are:**
    a. ask, listen, keeping safe
    b. keeping safe, ask, listen
    c. listen, keeping safe, ask
    d. none of the above

15. **It is important to move through a suicide intervention process while (circle all that apply):**
    a. keeping self and others safe by securing the platform
    b. ensuring trains return to service quickly
    c. communicating with transit control
    d. all of the above
    e. none of the above

16. **What will be your role in suicide intervention:**
    a. engage in counseling
    b. refer to appropriate resources
    c. call significant others
    d. insure the person gets on the next train
17. Helper attitude to someone at risk of suicide:
   a. should be made clear to a person at risk of suicide
   b. should be kept from person at risk of suicide
   c. may have either positive or negative effects
   d. have little to do with the process of intervention

18. Which of the following action plans would likely be most suitable for someone at risk of suicide:
   a. agreement for therapy
   b. agreement to meet again
   c. agreement to call a crisis line if troubled again with thoughts of suicide
   d. agreement to talk with significant other
   e. referral to appropriate resources

Part II

The following items represent a series of excerpts from intervention situations. Each excerpt begins with an expression by the patron concerning some aspect of the situation s/he faces, followed by two possible helper responses to the patron’s remark. You are to rate the response in term of how appropriate or inappropriate you feel the reply is to the patron’s comment. In the blank you should record a rating from -3 to +3, corresponding to the chart below. Be sure to respond to each item, and try not to leave any blanks.

+3 = Highly appropriate response
+2 = Appropriate response
+1 = Marginally appropriate response
0 = Neither appropriate nor inappropriate
-1 = Marginally inappropriate response
-2 = Inappropriate response
-3 = Highly inappropriate response

1. Patron 1:
   I decided to come down to the subway station tonight because I really feel things are too much for me……. Now my health is going downhill too, on top of all the rest. Without my husband around to care for me anymore, it just seems like the end of the world….
   
   _______ Helper A: Try not to worry too much about things. Everything will be all right.
   _______ Helper B: You must feel pretty lonely and afraid of what might happened… Are you thinking about suicide?

2. Patron 2:
   My life has been worthless since my wife, Emma, died four years ago. The kids are grown and married now, and I’ve been retired from my job at the railroad for some times. It just seems that I’d be better off dead.
   
   _______ Helper A: But try to think of what Emma would want for you. She would want you to continue leading a productive life, wouldn’t she?
   _______ Helper B: It sounds like everything just collapsed around you since
Emma died ... But what happened recently to make things even worse, to make you think dying is the only way out?

3. Patron 3:
How could you help me? Have you ever wanted to kill yourself?

______ Helper A: You sound like you are concern about whether I can understand and help you.

______ Helper B: Sure I have thought about suicide sometimes. But I also found more realistic solutions to my problems.

4. Patron 4:
I am so lonely so tired (crying). There just isn’t any where left to turn.

______ Helper A: You seem so alone, so miserable. Have you been feeling suicidal?

______ Helper B: Come on now. Things can’t be all that bad.

5. Patron 5:
Why should you care about me anyway?

______ Helper A: I’ve been trained to care about people. That’s my job.

______ Helper B: Because I think your death will be a terrible waste, and it concerns me that things are so that you are considering suicide. You need help to get through this critical period.

6. Patron 6:
No one can understand the kind of pain I’ve been going through. Sometimes I feel like I have to hurt myself..

______ Helper A: It seems that you’ve been suffering so much that hurting yourself is the only you can make the pain go away.

______ Helper B: But you are so young, you have so much to live for. How can you think of killing yourself?
APPENDIX H:

suicideAWARE three–months follow-up questionnaire
Dear TTC Employee:

The TTC and St Michael’s Hospital are working together to reduce the number of suicide attempts in the subway. We hope to be able to identify distressed people who might be thinking about suicide so that we can send help before they do. To achieve early identification of distressed at-risk TTC patrons, we have started, several months ago, a suicide prevention training program delivered to TTC Special Constables, Supervisors, and Instructors by the Trillium Health Centre.

The Suicide Studies Unit at St. Michael’s Hospital is evaluating the suicide prevention program and its impact on suicide knowledge, attitudes, and intervention skills. The program evaluation consists of completing 3 questionnaires, and as you may recall, you have completed 2 questionnaires, before and after the SuicideAWARE suicide prevention training workshop. The third similar follow-up questionnaire, measures the long-term impact of the training, and is completed 3 months after attending the SuicideAWARE workshop.

Attached to this letter is the follow-up questionnaire. Your participation is completely voluntary. The information you provide us is completely confidential, as is all the information you provided us at the time you’ve attended the workshop. The completed questionnaires are kept at St. Michael’s Hospital and are accessible only to the researchers in our unit. The TTC does not receive a report on any individual questionnaire. We will provide a group report to show how effective the overall training program was.

The completed questionnaire is mailed directly to the Arthur Sommer Rotenberg Chair in Suicide Studies at St. Michael’s Hospital. We have attached a self-addressed and pre-stamped envelop for your convenience. If you choose not to complete the questionnaire your employment status at the TTC would not be affected.

We thank you in advance for your cooperation and assistance in this important suicide prevention program evaluation.

Sincerely,

Dr Paul S. Links
Arthur Sommer Rotenberg Chair in Suicide Studies
St. Michael’s Hospital
My TTC Employee Number is: ___________________

Completion of this questionnaire is voluntary. If you do not wish to complete the questionnaire you do not have to answer the questions.
This is not a test but a survey of your opinions; there is no right or wrong answer, only your honest opinion counts. Please check one box:

1. Almost everyone at one time thought about suicide.
   □ Strongly agree    □ Agree     □ Undecided    □ Disagree    □ Strongly disagree

2. Most suicidal attempts are impulsive in nature.
   □ Strongly agree    □ Agree     □ Undecided    □ Disagree    □ Strongly disagree

3. Those who threaten suicide rarely do so.
   □ Strongly agree    □ Agree     □ Undecided    □ Disagree    □ Strongly disagree

4. Suicide is more prevalent among the very rich and the very poor.
   □ Strongly agree    □ Agree     □ Undecided    □ Disagree    □ Strongly disagree

5. Most people who try to kill themselves don’t really want to die.
   □ Strongly agree    □ Agree     □ Undecided    □ Disagree    □ Strongly disagree

6. Suicide happens without warning.
   □ Strongly agree    □ Agree     □ Undecided    □ Disagree    □ Strongly disagree

7. About 80% of those who commit suicide have attempted suicide at least once before.
   □ Strongly agree    □ Agree     □ Undecided    □ Disagree    □ Strongly disagree

8. It’s rare for someone who is thinking about suicide to be dissuaded by a ‘friendly ear.’
   □ Strongly agree    □ Agree     □ Undecided    □ Disagree    □ Strongly disagree

9. A large percentage of suicide victims come from a broken home.
   □ Strongly agree    □ Agree     □ Undecided    □ Disagree    □ Strongly disagree
10. The possibility of committing suicide is greater for older people (60 and over) than for young people (20 to 30 years).

□ Strongly agree □ Agree □ Undecided □ Disagree □ Strongly disagree

11. Once a person is suicidal, s/he is suicidal forever.

□ Strongly agree □ Agree □ Undecided □ Disagree □ Strongly disagree

12. Improvement following a suicidal crisis indicates that the risk is over.

□ Strongly agree □ Agree □ Undecided □ Disagree □ Strongly disagree

13. Once a person pursues a suicidal attempt, the probability of her/his trying again is minimal.

□ Strongly agree □ Agree □ Undecided □ Disagree □ Strongly disagree

14. Suicide attempters who use public places (such as a subway, a bridge) are more interested in getting attention.

□ Strongly agree □ Agree □ Undecided □ Disagree □ Strongly disagree

15. If someone wants to commit suicide, it is their business, and we should not interfere.

□ Strongly agree □ Agree □ Undecided □ Disagree □ Strongly disagree

16. A suicide attempt is essentially a “cry for help.”

□ Strongly agree □ Agree □ Undecided □ Disagree □ Strongly disagree

17. Usually, relatives of a suicide victim had no idea of what was about to happen.

□ Strongly agree □ Agree □ Undecided □ Disagree □ Strongly disagree

18. Those people who attempt suicide are usually trying to get sympathy from others.

□ Strongly agree □ Agree □ Undecided □ Disagree □ Strongly disagree

19. Potentially, every one of us can be a suicide victim.

□ Strongly agree □ Agree □ Undecided □ Disagree □ Strongly disagree

20. Most suicide victims are older persons with little to live for.

□ Strongly agree □ Agree □ Undecided □ Disagree □ Strongly disagree
TTC SUICIDE PREVENTION TRAINING PROGRAM

SUICIDE-RISK PROCEDURAL QUESTIONNAIRE

1. Which of the following is a good indicator that the passenger is at risk:
   a. They let several trains go by and they look depressed or distressed
   b. They are at the far end of the platform (train entrance point)
   c. They are agitated and pace at the far end of the platform close to the entrance to the tunnel
   d. They stand on the yellow line and they look depressed or distressed
   e. All of the above

2. Transit control should be called if a passenger on the platform:
   a. Extends part of their body beyond the platform
   b. Argues or fights with strangers
   c. Sits on the edge of the platform and dangle their feet over the rail
   d. Skateboards, rollerblades
   e. All of the above

3. Which of the following is a warning sign and help should be called:
   a. Passenger sits on the bench and cries
   b. Passenger takes off their clothes or removes their shoes
   c. Passenger says they are going to end their lives
   d. Passenger stands close to a foot bridge and appears depressed
   e. All of the above

4. Transit control should be called if a passenger:
   a. is inappropriately dressed for the weather
   b. is wearing army fatigues
   c. is at track level
   d. carries an umbrella on a sunny day
   e. none of the above
APPENDIX I:
safeTALK 3-months follow-up Questionnaire
Dear TTC Employee:

The TTC and St Michael’s Hospital are working together to reduce the number of suicide attempts in the subway. We hope to be able to identify distressed people who might be thinking about suicide so that we can send help before they do. To achieve early identification of distressed at-risk TTC patrons, we have started, several months ago, a suicide prevention training program delivered to TTC Special Constables, Supervisors, and Instructors by the Trillium Health Centre.

The Suicide Studies Unit at St. Michael’s Hospital is evaluating the suicide prevention program and its impact on suicide knowledge, attitudes, and intervention skills. The program evaluation consists of completing 3 questionnaires, and as you may recall, you have completed 2 questionnaires, before and after the safeTALK suicide prevention training workshop. The third similar follow-up questionnaire, measures the long-term impact of the training, and is completed 3 months after attending the safeTALK workshop.

Attached to this letter is the follow-up questionnaire. Your participation is completely voluntary. The information you provide us is completely confidential, as is all the information you provided us at the time you’ve attended the workshop. The completed questionnaires are kept at St. Michael’s Hospital and are accessible only to the researchers in our unit. The TTC does not receive a report on any individual questionnaire. We will provide a group report to show how effective the overall training program was.

The completed questionnaire is mailed directly to the Arthur Sommer Rotenberg Chair in Suicide Studies at St. Michael’s Hospital. We have attached a self-addressed and pre-stamped envelope for your convenience. If you choose not to complete the questionnaire your employment status at the TTC would not be affected.

We thank you in advance for your cooperation and assistance in this important suicide prevention program evaluation.

Sincerely,

Dr Paul S. Links
Arthur Sommer Rotenberg Chair in Suicide Studies
St. Michael’s Hospital
safeTALK
TTC SUICIDE PREVENTION TRAINING PROGRAM
Suicide Opinion Questionnaire
3 Months Follow-up

My TTC Employee Number is: ___________________

Completion of this questionnaire is voluntary. If you do not wish to complete the questionnaire you do not have to answer the questions. This is not a test but a survey of your opinions; there is no right or wrong answer, only your honest opinion counts. Please check one box:

1. Almost everyone at one time thought about suicide.
   - □ Strongly agree  □ Agree  □ Undecided  □ Disagree  □ Strongly disagree

2. Most suicidal attempts are impulsive in nature.
   - □ Strongly agree  □ Agree  □ Undecided  □ Disagree  □ Strongly disagree

3. Those who threaten suicide rarely do so.
   - □ Strongly agree  □ Agree  □ Undecided  □ Disagree  □ Strongly disagree

4. Suicide is more prevalent among the very rich and the very poor.
   - □ Strongly agree  □ Agree  □ Undecided  □ Disagree  □ Strongly disagree

5. Most people who try to kill themselves don’t really want to die.
   - □ Strongly agree  □ Agree  □ Undecided  □ Disagree  □ Strongly disagree

6. Suicide happens without warning.
   - □ Strongly agree  □ Agree  □ Undecided  □ Disagree  □ Strongly disagree

7. About 80% of those who commit suicide have attempted suicide at least once before.
   - □ Strongly agree  □ Agree  □ Undecided  □ Disagree  □ Strongly disagree

8. It’s rare for someone who is thinking about suicide to be dissuaded by a ‘friendly ear.’
   - □ Strongly agree  □ Agree  □ Undecided  □ Disagree  □ Strongly disagree

9. A large percentage of suicide victims come from a broken home.
   - □ Strongly agree  □ Agree  □ Undecided  □ Disagree  □ Strongly disagree
10. The possibility of committing suicide is greater for older people (60 and over) than for young people (20 to 30 years).

- Strongly agree
- Agree
- Undecided
- Disagree
- Strongly disagree

11. Once a person is suicidal, s/he is suicidal forever.

- Strongly agree
- Agree
- Undecided
- Disagree
- Strongly disagree

12. Improvement following a suicidal crisis indicates that the risk is over.

- Strongly agree
- Agree
- Undecided
- Disagree
- Strongly disagree

13. Once a person pursues a suicidal attempt, the probability of her/his trying again is minimal.

- Strongly agree
- Agree
- Undecided
- Disagree
- Strongly disagree

14. Suicide attempters who use public places (such as a subway, a bridge) are more interested in getting attention.

- Strongly agree
- Agree
- Undecided
- Disagree
- Strongly disagree

15. If someone wants to commit suicide, it is their business, and we should not interfere.

- Strongly agree
- Agree
- Undecided
- Disagree
- Strongly disagree

16. A suicide attempt is essentially a “cry for help.”

- Strongly agree
- Agree
- Undecided
- Disagree
- Strongly disagree

17. Usually, relatives of a suicide victim had no idea of what was about to happen.

- Strongly agree
- Agree
- Undecided
- Disagree
- Strongly disagree

18. Those people who attempt suicide are usually trying to get sympathy from others.

- Strongly agree
- Agree
- Undecided
- Disagree
- Strongly disagree

19. Potentially, every one of us can be a suicide victim.

- Strongly agree
- Agree
- Undecided
- Disagree
- Strongly disagree

20. Most suicide victims are older persons with little to live for.

- Strongly agree
- Agree
- Undecided
- Disagree
- Strongly disagree
TTC SUICIDE PREVENTION TRAINING PROGRAM

SUICIDE-RISK PROCEDURAL QUESTIONNAIRE

1. Which of the following is a good indicator that the passenger is at risk:

a. They let several trains go by and they look depressed or distressed

b. They are at the far end of the platform (train entrance point)

c. They are agitated and pace at the far end of the platform close to the entrance to the tunnel

d. They stand on the yellow line and they look depressed or distressed

e. All of the above

2. Transit control should be called if a passenger on the platform:

a. Extends part of their body beyond the platform

b. Argues or fights with strangers

c. Sits on the edge of the platform and dangle their feet over the rail

d. Skateboards, rollerblades

e. All of the above

3. Which of the following is a warning sign and help should be called:

a. Passenger sits on the bench and cries

b. Passenger takes off their clothes or removes their shoes

c. Passenger says they are going to end their lives

d. Passenger stands close to a foot bridge and appears depressed

e. All of the above

4. Transit control should be called if a passenger:

a. is inappropriately dressed for the weather

b. is wearing army fatigues

c. is at track level

d. carries an umbrella on a sunny day

e. none of the above
TTC SUICIDE PREVENTION TRAINING  
EVALUATION

Intervention Knowledge Test

Part I: Multiple Choice: For the questions which follow, select the best response by circling the appropriate letter. If uncertain, feel free to guess.

1. **Suicide is most likely a result of:**
   a. overwhelming stress  
   b. clinical depression  
   c. substance abuse  
   d. no single cause

2. **When confronted with the possibility of suicidal behaviour in a person, a helper should immediately:**
   a. refer the person to experienced suicide resources  
   b. discuss it directly with the person  
   c. call in significant others in the persons life  
   d. encourage the person to talk about positive aspects of his or her life

3. **Suicide plans are assessed on the basis of a persons:**
   a. level of planning  
   b. age  
   c. stated seriousness  
   d. apparent distress

4. **People who express suicide intentions:**
   a. clearly want to die  
   b. are ambivalent about dying  
   c. want to punish others  
   d. are manipulating

5. **Of the following, which provides the most important information in assessing the risk of suicide:**
   a. symptoms  
   b. stress  
   c. resources  
   d. physical health

6. **What generally determines if behaviour is suicidal?**
   a. the mood of the person  
   b. the intent of the person  
   c. the lethality of the method used  
   d. the history of prior behaviour

7. **Of the following, which provides the least important information in assessing the risk of suicide?**
   a. symptoms  
   b. stress  
   c. resources  
   d. physical health
8. Stress in people lives is:
   a. most commonly a subjective matter
   b. determined by external events
   c. synonymous with crisis
   d. highly related to level of suicide risk

9. If someone is showing signs and symptoms suggestive of suicide, a helper should:
   a. gather more information about what is bothering the person
   b. inquire about the support available from family and friends
   c. determine if substance abuse is a factor
   d. ask if the person is thinking about suicide

10. If someone admits to feeling suicidal, a helper should:
    a. calmly inquire about what is happening in their life
    b. find out if they’ve thought of how they would do it
    c. inform significant others
    d. arrange immediate referral

11. Which of the following phase comprise the safeTALK Model used in the Suicide Intervention Training Program:
    a. Invitation; tell, ask, listen, keeping safe
    b. Tell, invitation, listen, ask, keeping safe
    c. Ask, listen, invitations, tell, keep safe
    d. Keep safe, invitation, tell, ask, listen

12. Which of the following is considered an invitation for help:
    a. Actions
    b. Feelings
    c. Thoughts
    d. Reactions
    e. Life situations
    f. All of the above
    g. None of the above

13. Key tasks in the safeTALK Suicide Intervention Model are:
    a. ask, listen, keeping safe
    b. keeping safe, ask, listen
    c. listen, keeping safe, ask
    d. none of the above

14. It is important to move through a suicide intervention process while (circle all that apply):
    a. keeping self and others safe by securing the platform
    b. ensuring trains return to service quickly
    c. communicating with transit control
    d. all of the above
    e. none of the above

15. What will be your role in suicide intervention:
    a. engage in counseling
    b. refer to appropriate resources
    c. call significant others
    d. insure the person gets on the next train
16. **Helper attitude to someone at risk of suicide:**
   a. should be made clear to a person at risk of suicide
   b. should be kept from person at risk of suicide
   c. may have either positive or negative effects
   d. have little to do with the process of intervention

17. **Which of the following action plans would likely be most suitable for someone at risk of suicide:**
   a. agreement for therapy
   b. agreement to meet again
   c. agreement to call a crisis line if troubled again with thoughts of suicide
   d. agreement to talk with significant other
   f. referral to appropriate resources

**Part II**

The following items represent a series of excerpts from intervention situations. Each excerpt begins with an expression by the patron concerning some aspect of the situation s/he faces, followed by two possible helper responses to the patron’s remark. You are to rate the response in term of how appropriate or inappropriate you feel the reply is to the patron’s comment. In the blank you should record a rating from -3 to +3, corresponding to the chart below. Be sure to respond to each item, and try not to leave any blanks.

+3 = Highly appropriate response  
+2 = Appropriate response  
+1 = Marginally appropriate response  
0 = Neither appropriate nor inappropriate  
-1 = Marginally inappropriate response  
-2 = Inappropriate response  
-3 = Highly inappropriate response

1. **Patron 1:**
   I decided to come down to the subway station tonight because I really feel things are too much for me…… Now my health is going downhill too, on top of all the rest. Without my husband around to care for me anymore, it just seems like the end of the world….

   ________ Helper A: *Try not to worry too much about things. Everything will be all right.*

   ________ Helper B: *You must feel pretty lonely and afraid of what might happened… Are you thinking about suicide?*

2. **Patron 2:**
   My life has been worthless since my wife, Emma, died four years ago. The kids are grown and married now, and I’ve been retired from my job at the railroad for some times. It just seems that I’d be better off dead.

   ________ Helper A: *But try to think of what Emma would want for you. She would want you to continue leading a productive life, wouldn’t she?*

   ________ Helper B: *It sounds like everything just collapsed around you since*
Emma died ... But what happened recently to make things even worse, to make you think dying is the only way out?

3. Patron 3:
How could you help me? Have you ever wanted to kill yourself?

_______ Helper A: You sound like you are concerned about whether I can understand and help you.

_______ Helper B: Sure I have thought about suicide sometimes. But I also found more realistic solutions to my problems.

4. Patron 4:
I am so lonely so tired (crying). There just isn’t any where left to turn.

_______ Helper A: You seem so alone, so miserable. Have you been feeling suicidal?

_______ Helper B: Come on now. Things can’t be all that bad.

5. Patron 5:
Why should you care about me anyway?

_______ Helper A: I’ve been trained to care about people. That’s my job.

_______ Helper B: Because I think your death will be a terrible waste, and it concerns me that things are so that you are considering suicide. You need help to get through this critical period.

6. Patron 6:
No one can understand the kind of pain I’ve been going through. Sometimes I feel like I have to hurt myself.....

_______ Helper A: It seems that you’ve been suffering so much that hurting yourself is the only you can make the pain go away.

_______ Helper B: But you are so young, you have so much to live for. How can you think of killing yourself?
APPENDIX J:

Qualitative Interview Guides
TTC Subway Supervisors Interview Guide

Date: _______________

Audiotape Number: ______________

Overall instructions:

- Remind participant that the interview is confidential, voluntary and that s/he can refuse to answer any or all of the questions.
- Explain to the participant how the interview will be conducted.
- To encourage participants to answer questions fully, consider prompts such as:
  - Could you explain what you mean by that comment?
  - Could you give me an example of what you mean?

- Turn on the tape recorder.
DEMOGRAPHICS AND EMPLOYMENT

First, I would like to ask you some general questions about yourself and your employment at the TTC.

1. Your date of birth: ______

2. Sex: Male____ Female____

3. What is the highest level of education you have completed_______
   a. High-school ___
   b. High school___
   c. Some college or university___
   d. Completed college or university____

4. Your present marital status:
   a. Living with a partner
   b. Married_______
   c. Divorced or
   d. Separated______
   e. Widowed________
   f. Single, Never married____

5. How many years have you been employed at the TTC? ______

6. How many years have you been a Subway Supervisor? ______
   If less than total years employed at TTC:
   In what other capacity were you employed at the TTC?

7. What type of work did you do before you join the TTC?
   a. Management Occupations____
   b. Business, Finance and Administrative Occupations___
   c. Natural and Applied Sciences and Related occupations____
   d. Health Occupations
   e. Occupations in Social Sciences, Education, Government Service, and Religion
   f. Occupation in Arts, Culture, Recreation and Sports ___
   g. Sales and Service Occupations
   h. Trades, Transport and Equipment Operators and Related Occupations___
   i. Occupations unique to Primary Industry____
   j. Occupations Unique to Processing, Manufacturing and Utilities____
EXPERIENCE WITH SUICIDAL PATRONS

I would like to ask you some questions regarding your experience with suicide and/or suicidal patrons.

1. In all your years of employment at the TTC in how many completed suicides were you involved?
   # _____

2. In how many completed suicides were you involved in the past year? (Before safeTALK training)
   #______

3. In how many completed suicide were you involved in the past month (after safeTALK training)
   # ______

4. In all your years of employment at the TTC in how many attempted suicides were you involved?
   # ______

5. In how many attempted suicides were you involved in the past year? (Before safeTALK training)
   #______

6. In how many attempted suicides were you involved in the past month? (After the safeTALK training)
   #______

7. In your opinion, how many suicide attempts are there per year at the subway?

8. In your opinion, how many suicides are there each year at the subway?
   #_________

9. Why do you think people choose the subway to kill themselves?
TRAINING

Now I would like to ask you some questions regarding the suicide prevention training that was implemented at the TTC.

HAND OVER CARD.

1. To what extent are TTC employees aware of the suicide prevention training?

2. What aspect of the training did you find to be most relevant to your work?

3. What aspect of the training did you find to be most helpful to you in identifying patrons who might be suicidal?

4. What aspect of the training did you find least relevant to your work?

5. What aspect of the training did you find least helpful in identifying patrons who are at risk of suicide?

6. In your opinion, did the suicide prevention training change your views and attitude towards suicide?
   IF YES: How did your views and attitudes towards suicide change?
   IF NO: Why do you think the training did not change your view and attitudes towards suicide?

7. Do you think your attitude to suicide influences your alertness to patrons who might be at risk of suicide?
   IF YES: How does it influence your alertness?
   IF NO: How doesn’t your attitude influence your alertness/interaction with a suicidal patron?

8. Do you feel your alertness to suicide increased since the training?
   IF YES: In what way did the training contribute to the change in alertness?
   What aspect of the training contributed to the change in alertness?
9. After the training, how able are you at recognizing/identifying patrons at-risk of suicide?


10. After the training, how able are you at engaging a patron at-risk of suicide in direct and open talk about suicide?


11. Do you feel that the training helped increase your confidence level when dealing with suicidal patrons?

   IF YES: In what way did the training contribute to the change in confidence?
   What aspect of the training contributed to the change in confidence?

12. How competent do you feel you are at intervening with a suicidal patron?


13. How comfortable are you intervening with a suicidal patron?


14. Do you feel that the training helped you acquired new skills that improved your ability to intervene with patrons at risk of suicide?

   IF YES: Which new skills did you acquire?

   IF NO: What was lacking in the skill training?

15. Do you find that the procedures as outlined by the TTC hinder or help you intervene with a suicidal patron?

   IF HINDER: How does it hinder?

   IF HELP: How do the procedures help you when you intervene?

16. What else can the TTC do to prepare you to deal with a suicidal patron?
17. In your opinion, what else can be done to prevent suicides on the subway?

18. Do you have concerns regarding the suicide prevention program at the TTC?

19. Do you have any suggestions on how the suicide prevention program can be improved?

20. In your opinion, should the TTC have the ASIST trained resources available within the TTC?

PROMPT:
► What about implementing suicide prevention refresher workshops?
INTERVENTIONS

Now I would like to ask you some more specific questions about a recent intervention.

1. When was your most recent intervention?

2. Describe to me please your most recent suicide risk interventions.

2. Did the person alert you in some way to the possibility they were suicidal?
   □ Yes
   □ No

   **IF YES:** How did they alert you?

3. Did you use this information to help you ask openly and directly about suicide?
   □ Yes
   □ No

   **IF NO:** How did you find out they were suicidal?

4. Were you able to communicate to them that the seriousness of their concerns/the reasons for wanting to end their life are important to you?

   **IF YES:** How did you do this?

5. Did the intervention involve a MHA apprehension?
   □ Yes
   □ No

6. What aspect of the apprehension was the most difficult? What made it difficult?

7. Did this intervention occur prior to the suicide prevention training?
   □ Before
   □ After

   **IF YES:** Would you have handled it differently now that you have received the training? How?

   **IF NO:** If it was after the training, looking back, would you have handled the situation differently before the training?

8. How did the patron react to the intervention?
9. What was the patron’s attitude toward you at the beginning of the intervention?

10. Did the patron’s attitude toward you change by the end of the intervention?

   **IF YES:** In what way did the patron’s attitude changed?

   **IF NO:** What was the patron’s attitude throughout the intervention?

11. Was any other emergency service involved?

   **IF YES:** Which service was involved and what role did they play in the intervention?

12. Was the person referred to a community service?

   **IF YES:** By whom? To which community service were they referred?

   □ ER
   □ Gerstein
   □ Distress centre
   □ Mobile Crisis Team
   □ Other

   **IF NO:** Why not? What happened?

13. In your opinion, was the patron satisfied with the intervention and the referral?

   **IF YES:** What gave you that impression?

   **IF NOT:** How did the patron react?

14. How satisfied are you with how that particular intervention went?
STRESS AND COPING

Now I would like to ask you some questions about ways you cope with the suicide of a patron.

1. Have you had any experience with a suicide of a patron?

   **IF NO:** END OF INTERVIEW

2. What aspect of a suicide is the most traumatic or stressful to you?

3. If you could change any aspect of your role in the aftermath of suicide, what would you change?

4. Did you feel you were prepared or trained to deal with the suicide of a patron or your role in the aftermath of a suicide?

   **IF NOT:** How can the TTC prepare you better to deal with this aspect of your job?

5. What do you find is the most helpful to you after the suicide of a patron?

6. In your opinion, what is the least helpful to you after a completed or attempted suicide of a patron?

7. Is the debriefing offered after a suicide of a patron helpful to you?

   **IF YES:** What aspect is the most helpful?

   **IF NO:** Why not? What can be done differently?

8. Who do you feel comfortable talking with and sharing your feelings about a suicide of a patron?

   No one _____

   **IF NO ONE:** In general, do you find it difficult to share your feelings with others?

   (PROCEED TO QUESTION # 11)

   Spouse/partner _____ (PROCEED TO QUESTION # 11)

   Friend _____ (PROCEED TO QUESTION # 11)

   Counsellor/therapist _____ (PROCEED TO QUESTION # 9)
264

GP _____ (PROCEED TO QUESTION # 10)

Co-worker _____ (PROCEED TO QUESTION # 11)

9. Was the counsellor from the Employee and Family Assistance Program?

   IF NO: Why didn’t you use the Employee and Family Assistance Program? Did you use other Mental Health service in the community? Which community resource did you use?

10. Did you find it helpful to talk and share your feelings with a mental health professional/GP?

11. Do you rely on your co-workers for support?

   IF NOT: PROCEED TO QUESTION # 13

12. Is sharing your experience with a co-worker who has had a similar experience helpful?

   IF YES: Do you find it as helpful or more helpful than the debriefing?

13. Have you ever taken time-off because of the suicide of a patron?

   □ Yes

   IF YES: How long were you off work?

   □ No

14. Have you ever had unwanted memories (flashbacks or thoughts) of the suicide?

   IF NO: PROCEED TO QUESTION # 15

   IF YES: What were they like? How often had you had these Unwanted memories/ flashbacks? Did they last longer than a month? Have you had them in the past month?

15. Have you ever had any unpleasant dreams about the suicide of a patron?

   IF NO: What about being very upset when you were in a situation that reminded you of the suicide?

   IF NO: PROCEED TO QUESTION # 16
IF YES: Describe a typical dream. (What happens in them?)
How often had you had these unpleasant dreams?
Did they last longer than a month? Have you had any of these unpleasant dreams in the past month?

16. Have you ever sought help to deal with these unwanted memories (and/or unpleasant dreams)?

IF YES: What treatment did you received?
Did you have to take time off work? If yes, how long were you off work?

17. At a suicide scene, how do you deal with your emotions and carry on with your duties?

18. What can be done at the TTC to reduce your stress level when dealing with the aftermath of a suicide or a suicidal patron?

19. Are there any other supports that can be offered that in your opinion would be more helpful and more effective than the supports currently offered?

Is there anything else you would like to add? Was there anything that I forgot to ask in the interview that you feel it’s important for us to learn about?

Thank you for your cooperation and taking the time to assist us with the evaluation. Your contribution is extremely valuable to us.

GIVE OUT RESOURCES CARD.

Score Card


TTC Special Constables Interview Guide

Date: _______________

Audiotape Number: _______________

Overall instructions:

• Remind participant that the interview is confidential, voluntary and that s/he can refuse to answer any or all of the questions.

• Explain to the participant how the interview will be conducted.

• To encourage participants to answer questions fully, consider prompts such as:
  
  Could you explain what you mean by that comment?
  Could you give me an example of what you mean?

• Turn on the tape recorder.
Demographics and Employment

First, I would like to ask you some general questions about yourself and your employment at the TTC.

1. Your date of birth: ______

2. Sex: Male____ Female____

3. What is the highest level of education you have completed______
   k. High-school ___
   l. High school___
   m. Some college or university___
   n. Completed college or university____

4. Your present marital status:
   o. Living with a partner
   p. Married_______
   q. Divorced or
   r. Separated______
   s. Widowed_______
   t. Single, Never married____

5. How many years have you been employed at the TTC? ______

6. How many years have you been a Special Constable? _______

   If less than total years employed at TTC:
   In what other capacity were you employed at the TTC?

7. What type of work did you do before you join the TTC?
   u. Management Occupations____
   v. Business, Finance and Administrative Occupations____
   w. Natural and Applied Sciences and Related occupations____
   x. Health Occupations
   z. Occupation in Arts, Culture, Recreation and Sports____
   aa. Sales and Service Occupations
   bb. Trades, Transport and Equipment Operators and Related Occupations____
   cc. Occupations unique to Primary Industry____
   dd. Occupations Unique to Processing, Manufacturing and Utilities____
EXPERIENCE WITH MHA APPEHENSIONS

I would like to ask you some questions regarding your experience with Mental Health Act apprehensions.

1. In all your years of employment at the TTC in how many MHA apprehensions were you involved in?
   # _____

2. How many were suicide-risk apprehensions?
   # _____
   2.1 In how many MHA apprehensions were you involved in the past year? (Before safeTALK training)
      #_____
   2.2 How many were suicide-risk apprehensions?
      #_____
   2.3 In how many suicide interventions without MHA apprehension were you involved in the past month (after safeTALK training)
      #_____
   2.4 How many were suicide risk apprehensions?
      #_____
   2.5 How many suicide interventions without MHA apprehensions in the past month (after safeTALK training)?

3. In all your years of employment at the TTC in how many completed suicides were you directly involved
   #_____
   3.1 How many completed suicides in the past year? (Before safeTALK training)
      #_____
   3.2 How many completed suicides in the past month? (After the safeTALK training)
      #_____

4. In all your years of employment at the TTC in how many attempted suicides were you involved?
   #_____
   4.1 How many attempted suicides in the past year (before safeTALK training)?
      #_____

268
4.2 How many attempted suicides in the past month (after safeTALK training)?)
#______

5. In your opinion, how many MHA apprehensions are there per year at any TTC station, stop, property, or vehicle?
#________

6. In your opinion, how many suicide attempts are there per year at the subway?
#________

7. In your opinion, how many suicides are there each year at the subway?
#________

8. Why do you think people choose the subway to kill themselves?
TRAINING

Now I would like to ask you some questions regarding the suicide prevention training that was implemented at the TTC.

HAND OVER CARD.

1. To what extent are TTC employees aware of the suicide prevention training?


2. What aspect of the training did you find to be most relevant to your work?

3. What aspect of the training did you find to be most helpful to you in identifying patrons who might be suicidal?

4. What aspect of the training did you find least relevant to your work?

5. What aspect of the training did you find least helpful in identifying patrons who are at risk of suicide?

6. In your opinion, did the suicide prevention training change your views and attitude towards suicide?

   IF YES: How did your views and attitudes towards suicide change?

   IF NO: Why do you think the training did not change your view and attitudes towards suicide?

7. Do you think your attitude to suicide influences your alertness to patrons who might be at risk of suicide?

   IF YES: How does it influence your alertness?

   IF NO: How doesn’t your attitude influence your alertness/interaction with a suicidal patron?

8. Do you feel your alertness to suicide increased since the training?

   IF YES: In what way did the training contribute to the change in alertness?
What aspect of the training contributed to the change in alertness?

9. After the training, how able are you at recognizing/identifying patrons at-risk of suicide?


10. After the training, how able are you at engaging a patron at-risk of suicide in direct and open talk about suicide?


11. Do you feel that the training helped increase your confidence level when dealing with suicidal patrons?

    IF YES: In what way did the training contribute to the change in confidence?

    What aspect of the training contributed to the change in confidence?

12. How competent do you feel you are at intervening with a suicidal patron?


13. How comfortable are you intervening with a suicidal patron?


14. Do you feel that the training helped you acquired new skills that improved your ability to intervene with patrons at risk of suicide?

    IF YES: Which new skills did you acquire?

    IF NO: What was lacking in the skill training?

15. Do you find that the procedures as outlined by the TTC hinder or help you intervene with a suicidal patron?

    IF HINDER: How does it hinder?

    IF HELP: How do the procedures help you when you intervene?
16. What else can the TTC do to prepare you to deal with a suicidal patron?

17. In your opinion, what else can be done to prevent suicides on the subway?

18. Do you have concerns regarding the suicide prevention program at the TTC?

19. Do you have any suggestions on how the suicide prevention program can be improved?

20. In your opinion, should the TTC have the ASIST trained resources available within the TTC?

PROMPT:
➤What about implementing suicide prevention refresher workshops?
Now I would like to ask you some more specific questions about a recent apprehension/intervention.

1. When was your most recent MHA apprehension?
2. Describe to me please, your most recent MHA apprehensions/suicide risk interventions.
3. Did the person alert you in some way to the possibility they were suicidal?
   - Yes
   - No
   **IF YES:** How did they alert you?
4. Did you use this information to help you ask openly and directly about suicide?
   - Yes
   - No
   **IF NO:** How did you find out they were suicidal?
5. Were you able to communicate to them that the seriousness of their concerns/the reasons for wanting to end their life are important to you?
   **IF YES:** How did you do this?
6. Did the intervention involve a MHA apprehension?
   - Yes
   - No
7. What aspect of the apprehension was the most difficult? What made it difficult?
8. Did this apprehension/intervention occur prior to the suicide prevention training?
   - Before
   - After
   **IF YES:** Would you have handled it differently now that you have received the training? How?
   **IF NO:** If it was after the training, looking back, would you have handled the situation differently before the training?
9. How did the patron react to the apprehension/intervention?
10. What was the patron’s attitude toward you at the beginning of the intervention?
11. Did the patron’s attitude toward you change by the end of the intervention?

   **IF YES:** In what way did the patron’s attitude changed?

   **IF NO:** What was the patron’s attitude throughout the intervention?

12. Was any other emergency service involved?

   **IF YES:** Which service was involved and what role did they play in the apprehension/intervention?

13. Was the person referred to a community service?

   **IF YES:** By whom? To which community service were they referred?

   □ ER
   □ Gerstein
   □ Distress centre
   □ Mobile Crisis Team
   □ Other

   **IF NO:** Why not? What happened?

14. In your opinion, was the patron satisfied with the intervention and the referral?

   **IF YES:** What gave you that impression?

   **IF NOT:** How did the patron react?

15. How satisfied are you with how that particular intervention went?
STRESS AND COPING

Now I would like to ask you some questions about ways you cope with the suicide of a patron.

1. Have you had any experience with a suicide of a patron?

   **IF NO: END OF INTERVIEW**

2. What aspect of a suicide is the most traumatic or stressful to you?

3. If you could change any aspect of your role in the aftermath of suicide, what would you change?

4. Did you feel you were prepared or trained to deal with the suicide of a patron or your role in the aftermath of a suicide?

   **IF NOT:** How can the TTC prepare you better to deal with this aspect of your job?

5. What do you find is the most helpful to you after the suicide of a patron?

6. In your opinion, what is the least helpful to you after a completed or attempted suicide of a patron?

7. Is the debriefing offered after a suicide of a patron helpful to you?

   **IF YES:** What aspect is the most helpful?

   **IF NO:** Why not? What can be done differently?

8. Who do you feel comfortable talking with and sharing your feelings about a suicide of a patron?

   No one _____

   **IF NO ONE:** In general, do you find it difficult to share your feelings with others?

   (PROCEED TO QUESTION# 11)

   Spouse/partner _____ (PROCEED TO QUESTION # 11)

   Friend _____ (PROCEED TO QUESTION # 11)

   Counsellor/therapist _____ (PROCEED TO QUESTION # 9)
GP _____ (PROCEED TO QUESTION # 10)

Co-worker _____ (PROCEED TO QUESTION # 11)

9. Was the counsellor from the Employee and Family Assistance Program?

   IF NO: Why didn’t you use the Employee and Family Assistance Program? Did you use other Mental Health service in the community? Which community resource did you use?

10. Did you find it helpful to talk and share your feelings with a mental health professional/GP?

11. Do you rely on your co-workers for support?

   IF NOT: PROCEED TO QUESTION # 13

12. Is sharing your experience with a co-worker who has had a similar experience helpful?

   IF YES: Do you find it as helpful or more helpful than the debriefing?

13. Have you ever taken time-off because of the suicide of a patron?
   □ Yes     
   IF YES: How long were you off work?
   □ No

14. Have you ever had unwanted memories (flashbacks or thoughts) of the suicide?

   IF NO: PROCEED TO QUESTION # 15

   IF YES: What were they like? How often had you had these Unwanted memories/flashbacks? Did they last longer than a month? Have you had them in the past month?

15. Have you ever had any unpleasant dreams about the suicide of a patron?

   IF NO: What about being very upset when you were in a situation that reminded you of the suicide?

   IF NO: PROCEED TO QUESTION # 16

   IF YES: Describe a typical dream. (What happens in them?)
How often had you had these unpleasant dreams? Did they last longer than a month? Have you had any of these unpleasant dreams in the past month?

16. Have you ever sought help to deal with these unwanted memories (and/or unpleasant dreams)?

**IF YES:** What treatment did you received? Did you have to take time off work? If yes, how long were you off work?

17. At a suicide scene, how do you deal with your emotions and carry on with your duties?

18. What can be done at the TTC to reduce your stress level when dealing with the aftermath of a suicide or a suicidal patron?

19. Are there any other supports that can be offered that in your opinion would be more helpful and more effective than the supports currently offered?

Is there anything else you would like to add? Was there anything that I forgot to ask in the interview that you feel it’s important for us to learn about?

Thank you for your cooperation and taking the time to assist us with the evaluation. Your contribution is extremely valuable to us.

**GIVE OUT RESOURCES CARD.**

**Score Card**


APPENDIX K:
Effect size formulas for repeated measures
ES = \frac{\text{Mean}_{\text{post}} - \text{Mean}_{\text{pre}}}{\text{Pooled variance}} \times \sqrt{1 - \text{Correlation}}

ES = \frac{\text{Mean}_{\text{post}} - \text{Mean}_{\text{pre}}}{(N_2 - 1) \times SD_1 + (N_2 - 1) \times (N_1 + N_2 - 2)} \times \sqrt{1 - r}
APPENDIX L:
Consent forms
Consent to Participate in a Research Study
Information Form

You are being invited to participate in a research study to help us evaluate the suicide prevention program implemented at the TTC.

Before agreeing to participate in this research study, it is important that you read and understand this research consent form. This form provides all the information we think you will need to know in order to decide whether you wish to participate in the study. If you have any questions after you read this form, please ask the researcher any questions that you may have. You should not sign the form until you are sure that you understand everything on this form.

Title of Research Study
Key Informants Survey of Suicide Warning Signs: A Collaborative Project between the Arthur Sommer Rotenberg Chair in Suicide Studies and Toronto Transit Commission (TTC)

Investigator(s):
Paul S. Links, MD FRCP (C)
Arthur Sommer Rotenberg Chair in Suicide Studies
St. Michael’s Hospital
30 Bond Street,
Room 2-010, Shuter Wing
Toronto, ON M5B 1W8
Telephone number: (416) 864-6060 ext. 2689
Study Sponsor
Toronto Transit Commission.

Purpose of the Research:
St. Michael’s Hospital and the TTC are working together in evaluating the suicide prevention training program implemented at the TTC. The purpose of the program evaluation is to receive feedback from employees who had attended the “safeTALK” training session. Specifically, we would like to evaluate the impact the training had on your ability to identify distress at risk of suicide TTC patrons and your comfort and confidence levels when you interact with them.

Description of the Research:
TTC senior management and your Union Representative asked if you are willing to participate in this study and be interviewed by the research coordinator. You were approached because you had attended

the “safeTALK” training session given by Trillium Health Centre. TTC management and your Union Representative felt that you can help us with our program evaluation.

Before the interview the research coordinator will explain the interview to you and go over this consent form. You will be asked if you have any questions regarding the interview. If you agree to participate you will need to sign the consent form. Only after you signed the consent form will the research coordinator interview you. In the interview, the research coordinator will ask you questions about your experiences with suicidal TTC patrons, your opinion about the suicide prevention training and the effect it had on your ability to identify and help patrons in distress. You will also be given the opportunity to add any other comments you may have.

The interview will take about 90 minutes and will be tape-recorded. The recorded audio tape is destroyed after the interview has been typed and the typing verified.

Potential Harms (Injury, Discomforts or Inconvenience):
There are no known harms associated with participation in this study. However, you will be asked some questions that might make you feel somewhat uncomfortable and remind you of the suicide(s) you witnessed. If you do feel uncomfortable, you may refuse to answer any question, or terminate your participation in this study at any point in time, without any adverse consequences or without fear of reprisal to your health care and employment status.

It is possible that in the past you have had clinical contact with some of the study investigators. You may have clinical contact with them in the future. However, the information you provide during the interviews will not be added to your medical record nor will the information you provide impact on the health care that you receive. You should know that there is an important exception concerning the acute risk of harm to yourself or others, which is discussed below under Confidentiality and Privacy.

Potential Benefits:
You will not benefit directly from participating in this study. The information you provide will help us evaluate and enhance of the suicide prevention training program for TTC employees.
**Confidentiality and Privacy:**
Confidentiality will be respected and no information that discloses your identity will be released or published without consent, unless required by law.

During the interview, if the researcher is seriously concerned that you may harm yourself or someone else in the very near future, the researcher is required by law to report these concerns to the proper authorities and enlisting appropriate medical assistance, such as your family doctor, or appropriate mental healthcare staff. These health care staff would then decide on what step to take in order to ensure that you or someone else remains safe.

If, at any time during the study, you wish to speak with a doctor, nurse, or counsellor, the researcher has a list of people you can call and/or the researcher will arrange for you to speak with someone, as necessary.

**Publication of results:**
Results of this study will be reported in professional and scientific publications and at conferences. The information reported from this study wills not identity you in any way.

**Reimbursement:**
Your time will be reimbursed by the TTC.

**Participation and Withdrawal:**
Participation in research is voluntary. If you choose not to participate, your employment with the TTC will not be affected. You will also have access to customary care at SMH.

If you choose to participate in this study you can withdraw from the study at any time without any effect on your employment with the TTC or the care you will receive at SMH.

Withdrawal from the study does not necessarily include withdrawal of any data compile up to that point.

**Research Ethics Board Contact:**
If you have any questions about your rights as a research participant you may contact Dr. Julie Spence, Chair, St. Michael’s Hospital, Research Ethics Board at (416) 864-6060 Ext 2557.
Consent to participate in a research Study
Consent Form

I acknowledge that the research study described above have been explained to me and that any questions that I have asked have been answered to my satisfaction. I have been informed of the alternatives to participation in this study, including the right not to participate and the right to withdraw without compromising the quality of medical care at St. Michael’s Hospital for me and for other members of my family nor will it affect my employment status at the TTC. As well, the potential risks, harms and discomforts have been explained to me and I also understand the benefits (if any) of participating in the research study.

I understand that I have not waived my legal rights nor released the investigators, sponsors, or involved institutions from their legal and professional duties. I know that I may ask now, or in the future, any questions I have about the study or the research procedures. I have been assured that records relating to me and my care will be kept confidential and that no information will be released or printed that would disclose personal identity without my permission unless required by law. I have been given sufficient time to read and understand the above information.

I hereby consent to participate, and will be given a copy of this consent form.

___________________________________                        ______________________
Name of Participant                 Signature of participant

____________________________________           _____________________________
Name of Research Coordinator  Signature of Research Coordinator
APPENDIX M:
List of Community Resources
List of Community Resources

If you feel distressed you can call:

**Employee and Family Assistance Program**: 1-800-572-0039

**Distress Centres of Toronto**: (416) 408-4357; (416) 408-HELP

**The Gerstein Centre**: (416) 929-5200

**Telecare Distress Centre Etobicoke**: (416) 247-5426

**Distress Centre Oakville**: (905) 849-4541

**Distress Centre Peel**: (905) 278-7208

**Telecare Distress Centre Brampton**: (905) 459-7777

The above crisis lines are open 24 hours 7 days a week

**OR**

You may go to St. Michael’s Hospital Emergency Department or the Emergency Department at a Hospital nearest to your home

**OR**

You may contact your family physician.
APPENDIX N:
Linked and Unlinked Tables
Table A: Summary of Linked and Unlinked fixed effects

<table>
<thead>
<tr>
<th>Measure</th>
<th>Unlinked Data</th>
<th>Linked Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPQ</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Period</td>
<td>$F_{2,195}=14.086$ p=.001</td>
<td>$F_{2,243}=6.946$ p=.001</td>
</tr>
<tr>
<td></td>
<td>Pre-Post 0.001</td>
<td>Pre-Post 0.008</td>
</tr>
<tr>
<td></td>
<td>Pre-FU 0.147</td>
<td>Pre-FU 0.002</td>
</tr>
<tr>
<td>Period X job group</td>
<td>$F_{6,242}=3.177$ p=.005</td>
<td>$F_{6,194}=2.155$ p=.049</td>
</tr>
<tr>
<td>SOQ-R</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Period</td>
<td>$F_{2,214}=8.967$ p=.001</td>
<td>$F_{2,238}=11.384$ p=.001</td>
</tr>
<tr>
<td></td>
<td>Pre-Post 0.001</td>
<td>Pre-Post 0.001</td>
</tr>
<tr>
<td></td>
<td>Pre-FU 0.237</td>
<td>Pre-FU 0.001</td>
</tr>
<tr>
<td>Period X job group</td>
<td>$F_{6,211}=2.393$ p=.029</td>
<td>$F_{6,218}=6.11$ p=.722</td>
</tr>
<tr>
<td>IKT-R</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Period</td>
<td>$F_{2,83}=31.269$ p=.001</td>
<td>$F_{2,118}=19.277$ p=.001</td>
</tr>
<tr>
<td></td>
<td>Pre-Post 0.001</td>
<td>Pre-Post 0.001</td>
</tr>
<tr>
<td></td>
<td>Pre-FU 0.008</td>
<td>Pre-FU 0.001</td>
</tr>
<tr>
<td>Period X job group</td>
<td>$F_{4,83}=.578$ p=.679</td>
<td>$F_{4,117}=.234$ p=.919</td>
</tr>
<tr>
<td>SIRI-R</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Period</td>
<td>$F_{2,65}=8.335$ p=.001</td>
<td>$F_{2,116}=20.359$ p=.001</td>
</tr>
<tr>
<td></td>
<td>Pre-Post 0.001</td>
<td>Pre-Post 0.001</td>
</tr>
<tr>
<td></td>
<td>Pre-FU 0.330</td>
<td>Pre-FU 0.001</td>
</tr>
<tr>
<td>Period X job group</td>
<td>$F_{4,65}=.482$ p=.749</td>
<td>$F_{4,116}=1.467$ p=.217</td>
</tr>
</tbody>
</table>
Table B: SOQ-R Mean Scores of Linked and Unlinked Data

<table>
<thead>
<tr>
<th>Type of Data</th>
<th>SOQ-R Data</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td></td>
</tr>
<tr>
<td>Linked</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-training</td>
<td>11.26</td>
<td>2.96</td>
<td></td>
</tr>
<tr>
<td>Post-training</td>
<td>12.61</td>
<td>2.98</td>
<td></td>
</tr>
<tr>
<td>Follow-up</td>
<td>12.35</td>
<td>2.76</td>
<td></td>
</tr>
<tr>
<td>Unlinked</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-training</td>
<td>10.69</td>
<td>2.85</td>
<td></td>
</tr>
<tr>
<td>Post-training</td>
<td>12.14</td>
<td>3.31</td>
<td></td>
</tr>
<tr>
<td>Follow-up</td>
<td>11.25</td>
<td>2.41</td>
<td></td>
</tr>
</tbody>
</table>
Table C: Immediate and Long-term Effects of Training on Linked Raw Mean Scores of Outcome Measures

<table>
<thead>
<tr>
<th>Measure</th>
<th>Linked Raw Mean Scores</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Mean</td>
<td>SD</td>
<td>P value</td>
</tr>
<tr>
<td>SOQ-R</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-training</td>
<td>127</td>
<td>11.26</td>
<td>2.96</td>
<td></td>
</tr>
<tr>
<td>Post-training</td>
<td>127</td>
<td>12.61</td>
<td></td>
<td>&lt; 0.001&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Follow-up</td>
<td>127</td>
<td>12.35</td>
<td>2.76</td>
<td>&lt; 0.001&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>SPQ</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-training</td>
<td>127</td>
<td>3.65</td>
<td>0.64</td>
<td></td>
</tr>
<tr>
<td>Post-training</td>
<td>127</td>
<td>3.77</td>
<td></td>
<td>&lt; 0.001&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Follow-up</td>
<td>127</td>
<td>3.89</td>
<td>0.54</td>
<td>&lt; 0.001&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>IKT-R</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-training</td>
<td>63</td>
<td>7.64</td>
<td>2.08</td>
<td></td>
</tr>
<tr>
<td>Post-training</td>
<td>63</td>
<td>9.64</td>
<td></td>
<td>&lt; 0.001&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Follow-up</td>
<td>63</td>
<td>8.95</td>
<td>1.73</td>
<td>&lt; 0.001&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>SIRI-R</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-training</td>
<td>63</td>
<td>6.98</td>
<td>2.62</td>
<td></td>
</tr>
<tr>
<td>Post-training</td>
<td>63</td>
<td>8.35</td>
<td></td>
<td>&lt; 0.001&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Follow-up</td>
<td>63</td>
<td>8.80</td>
<td>1.94</td>
<td>&lt; 0.001&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

<sup>a</sup> = Pre-post comparisons

<sup>b</sup> = Pre-follow-up comparisons

<sup>c</sup> = Based on Cohen’s d for repeated measures