

NOTE TO USERS

This reproduction is the best copy available.

UMI[®]

A Systematic Review of Instructional Interventions to Improve School Completion:
Mapping the Evidence

Lori Mae Wozney

A Thesis
In the Department
of
Education

Presented in Partial Fulfilment of the Requirements
For the Degree of Doctorate in Philosophy (Educational Technology)
at

Concordia University
Montreal, Quebec, Canada

January 2009

©Lori Mae Wozney, 2009



Library and Archives
Canada

Published Heritage
Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et
Archives Canada

Direction du
Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file *Votre référence*
ISBN: 978-0-494-63418-9
Our file *Notre référence*
ISBN: 978-0-494-63418-9

NOTICE:

The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

■ ♦ ■
Canada

ABSTRACT

Systematic Review of Instructional Interventions to Improve School Completion: Mapping the Evidence

Lori Wozney, Ph.D.
Concordia University, 2008

This review of research on dropout prevention programs in Canada between 1990 and 2006 was conducted with the goal of increasing awareness and knowledge of (a) current research on high school dropout prevention and intervention in Canada; (b) instructional design and implementation of successful programs; (c) context-related factors that moderate program effectiveness and (d) selecting and/or designing programs that take into consideration current research evidence.

The identification of studies to be used in this review was conducted through a comprehensive search of publicly available literature (i.e., research databases, contacting researchers and program administrators, contacting local and provincial education agencies, etc.). Of the 240 documents retrieved 38 met all of the inclusion criteria. An additional 30 studies from outside of Canada were also analyzed.

Underreporting and missing data presented significant challenges in terms of analyzing instructional practices and impacts. Results showed that the most frequent type of dropout prevention programs were pull-out support, specialized courses and workshops or alternative schools. Most programs incorporated multiple forms of support (i.e., combinations of health services, life-skills, career preparation, academic support, cultural/spiritual enrichment, etc.). Instructional strategies varied across findings with the most common being tutoring, work-based learning and mentoring.

Almost three quarters of the findings presented evidence of positive program outcomes with another 13% reporting strong positive program impacts.

Future research might focus on linking outcome impacts (e.g., enrolment status/dropout rates) with the program performance context to look beyond “the learner” as the site of dropout prevention. Application for stakeholders and practitioners includes, among others, recommendations for revisiting existing practices and policies to determine if mainstream classroom practices support the school/work connection and redefining instruction using best-practices in teaching to accommodate self-direction, flexibility in course delivery and responsiveness to the needs of at-risk learners.

I am incredibly thankful for the support of my husband, Paul who kept faith in me long after mine had started to falter; and my children Jonah and Lois for being far more patient and generous than anyone could ask of their children (your unique snack creations during the writing process were also appreciated).

Table of Contents

List of Tables	ix
Executive Summary	1
Current Canadian Issues in Dropout Prevention Programming	7
Canadian Historical Context	7
Identifying Underlying Factors Relating to Dropout.....	15
Defining Educational Programming for Dropout Prevention.....	19
<i>Early Prevention Programs</i>	19
<i>Intervention Programs (Junior And Senior)</i>	20
<i>School-to-Work Transition Strategies</i>	22
<i>Re-entry/Retrieval Initiatives</i>	23
Characteristics of Effective Prevention Programs	23
The Role of HPT in Public Education Interventions	29
Politics and Purpose: HPT in Traditional Educational Contexts	31
Accountability vs. Autonomy in Designing Educational Programming ...	33
Program Design and Delivery: Fidelity to Implementation Plans	34
HPT and the Use of Systematic Reviews.....	36
Purpose & Research Questions.....	40
Methodology	42
Systematic Reviews.....	42
The Value of Systematic Review for Dropout Interventions	47
<i>Making the Unmanageable Manageable</i>	47
<i>Establishing Consistency of Findings</i>	47
<i>Integrating Critical Information For Decision-Making</i>	48
<i>Comparing the Influence of Research Design</i>	48
<i>Encouraging Theory Building</i>	49

Limitations and Challenges in Conducting Systematic Reviews	49
<i>The Quality and Interpretations of Primary Studies</i>	50
<i>Lack of Common Language and Reporting Styles</i>	51
<i>Achieving Valid Coding of Primary Studies</i>	53
<i>Political Challenges</i>	54
Locating Available Program Documentation	54
<i>Identifying Relevant Search Terminology</i>	55
<i>Developing Search Strategies</i>	57
Inclusion /Exclusion Criteria	60
Developing Coding Scheme	63
Quantitative Methods for Extracting Data.....	66
Qualitative Methods for Extracting Data	68
Assessing Research Quality.....	69
Results	71
Outcome of Searches	71
Coding of Studies	75
Analysis of Effect Sizes	77
Demographic Features	86
<i>Research Design</i>	86
<i>Selection Criteria for Program/Intervention</i>	87
<i>Participant Characteristics</i>	88
Program Features.....	88
<i>School Level</i>	88
<i>Program Type</i>	89
<i>Duration of Intervention</i>	92
<i>Program Components</i>	92
<i>Instructional Design</i>	94
<i>Support Staff</i>	96
<i>Implementation Considerations</i>	97

Program Effectiveness.....	99
<i>Outcome Categories</i>	99
<i>Strength of Outcome Evidence</i>	100
Summary and Discussion	103
How are Programs Measuring “Effectiveness”	103
Methodological Quality of Current Research.....	108
How Have Programs Been Designed and Delivered?	112
Application for Practitioners	117
Rethinking the Existing System	118
Routinely Examining and Reporting on Relevant Data	121
Conclusion	125
References	127
APPENDICES	136
Appendix A- Search Strategy Summary.....	136
Appendix B- Coding Scheme	141
Appendix C- Effect Size Conversion Formulas	149
Appendix D- Sample Form Letters	151
Appendix E- Summary of Effect Size Extractions	152
Appendix F- Summary of Effect Sizes by Country of Study Origin	155
Appendix G- List of Included Studies/Reports.....	158

List of Tables

<i>Table 1. Key Characteristics of Dropout Prevention/School Completion Programs</i>	25
<i>Table 2. Types of Secondary Research Reviews</i>	44
<i>Table 3. Phases in Conducting a Systematic Review (EPPI guidelines)</i>	46
<i>Table 4. Guidelines for Systematic Reviews</i>	64
<i>Table 5. Criteria for Assessing Research Quality</i>	71
<i>Table 6. Categories, Numbers and Percentages of Excluded Canadian Studies</i>	75
<i>Table 7. Publication Information for Sixty-Eight Written Reports</i>	76
<i>Table 8. Weighted Mean Effect Sizes for Combined Outcome Categories</i>	81
<i>Table 9. Weighted Mean Effect Sizes Based on Program Complexity</i>	83
<i>Table 10. Weighted Mean Effect Sizes on Physical Presence Outcomes for Canadian and Non-Canadian Programs</i>	84
<i>Table 11. Research Designs Compared Between Canadian and Non-Canadian Based Programs</i>	86
<i>Table 12. Selection Criteria for Canadian and Non-Canadian Based Programs</i>	87
<i>Table 13. Comparison of Targeted Grade Level</i>	89
<i>Table 14. Reported Model and Schedule Designs for All Programs</i>	90
<i>Table 15. Comparison of Program Models for Canadian and Non-Canadian Based- Programs</i>	91
<i>Table 16. Program Components for Canadian and Non-Canadian Based Programs</i>	93
<i>Table 17. Summary of Program Components and Instructional Strategies Across All Interventions</i>	95
<i>Table 18. Instructional Strategies Used by Canadian and Non-Canadian Based Programs</i>	96
<i>Table 19. Qualitative Descriptors of Program Value and Barriers</i>	98
<i>Table 20. Categories of Program Objectives for Total Sample (n=151)</i>	100
<i>Table 21. Summary of Effectiveness Evidence for Total Sample (n=151)</i>	101
<i>Table 22. Summary of Instructional Strategies by Outcome Strength</i>	102
<i>Table 23. Summary of Validity Ratings Compared by Outcome Strength</i>	103

Executive Summary

The number of students not completing high school is receiving increased attention as a serious challenge facing the North American educational system (e.g., Barton, 2005; King, 2005). Stats Canada reported that during the 2004/2005 academic year 1 in 10 Canadian students dropped out of high school, and while the overall dropout rate in Canada has shown significant improvement in the last decade dropout rates remain disproportionately high among young Canadian males and youth living in rural communities (King, 2005; Bowlby, 2005). Quebec had one of the nations highest dropout rates in the 2004/2005 year at just under 13% with almost double the number of rural students dropping out (19.5%) compared to their urban counterparts. While some are quick to point to the effectiveness of dropout prevention programs as the reason for declining dropout rates (e.g., Bowlby, 2005; Canadian Council on Learning, 2005) little is actually known about these programs and the kinds of instructional practices they are employing. In fact, dropout rates in provinces like Alberta are among the highest in the country despite having increased education spending and “one of the most successful education systems in the world” (Little, p.1).

This review of research on dropout prevention programs in Canada between 1990 and 2006 was conducted with the goal of increasing awareness and knowledge of (a) current research on high school dropout prevention and intervention in Canada; (b) instructional techniques and program features of successful programs; (c) context-related factors that may moderate program effectiveness and (d) how to select and/or design programs for high school dropout prevention that take into consideration current research evidence.

The identification of available qualitative and quantitative studies was conducted through a comprehensive search of publicly available literature (i.e., research databases, contacting researchers and program administrators, contacting local and provincial education agencies, etc.). Of the 240 documents retrieved for possible inclusion 38 met all of the inclusion criteria. An additional 30 studies from outside of Canada were also analyzed for comparison.

Each program was coded to extract data relating to program design, implementation and outcomes. Program features explored fell into the following categories: study identification (e.g., author, year, source of publication); methodological features (e.g., research design, effect sizes, durability of treatment, sample size, etc.); objectives (e.g., program objectives, student population targeted, etc.); environmental features (e.g., urban/rural, age group, linguistic or racial composition of students, etc.); resources (e.g., program funding source, support personnel, costs benefit, training, etc.); components (e.g., use of technology, degree of self-paced instruction, student/teacher ratio, types of assessment, degree of mentoring or peer support, etc.); and management (e.g., use of external evaluator, type of program, duration of intervention, etc.)

Underreporting and missing data presented significant challenges in terms of analyzing program practices and impacts. This has been reported in other reviews of this literature (e.g. Prevatt & Kelly, 2003; Lehr et al, 2003). Many of the studies did not report specifically on “dropout rates” as a dependent variable but targeted other kinds of program outcomes. A five category model presented in Lehr et al (2003) was used to cluster the indicators of program effectiveness into: Academic/cognitive = GPA, standardized test scores, study habits; Physical presence = enrolment status, attendance,

graduation rate, dropout rate; Psychological = student attitudes towards learning or school, self-esteem, depression; Social Behavioural = problem behaviour, social competence, drug use, violent behaviour; and Support for learning = student attitude toward teachers, school climate. The majority of programs used measures of “physical presence” as the key indicator of program effectiveness with psychological outcomes like attitudes toward learning and school, self –esteem being the second most common group of indicators.

It is important to note that dropout rates and attendance measures were calculated across programs using a variety of formulas. Some programs defined dropout rates as the proportion of students in a given age range who leave school each year without completing a high school program. Some programs provide cumulative data on dropouts among all, for example, 16- through 24-years-olds who are not enrolled in school and who have not earned a high school diploma or its equivalent. Other programs measured dropout rates as cohorts; following an identifiable group of students over time and tracking their progress through the system. This disparity in the way outcomes were reported raises interesting issues about how each formula provides different kinds of information about that population of students and how we need to think about using student self-reported data, the accuracy of census survey data and difficulties in tracking traditionally highly mobile groups of students, (e.g., in urban areas where students can move from school to school within the same board) might affect the way we interpret their data if it is being used to make instructional design and program decisions.

The methodological quality of available Canadian studies was generally poor, lacking in both descriptive data and statistical analysis of program impacts. This finding was

particularly relevant given emerging regional efforts (e.g., the NANS initiative in Quebec and the Imagine Our Schools initiative in Nova Scotia) to implement rigorous assessment procedures within educational programs. Martine Durand, deputy director of the Directorate for Employment, Labour and Social affairs recently argued that “there should be more accountability on the part of schools” and that the lack of government involvement, particularly for Aboriginal youth is “one of the weak spots in Canada’s education system” (McKibbin, 2008).

Results of our review showed that the most frequent type of dropout prevention programs were pull-out support, specialized courses and workshops style interventions. Most programs incorporated multiple forms of support (i.e., combinations of health services, life-skills, career preparation, academic support, cultural/spiritual enrichment, parental development and vocational training) with academic support, life skills (e.g., personal care, behaviour modification, budgeting, time management) and vocational training having the strongest focus of programs reviewed. Analysis of effect sizes for studies measuring dropout and attendance rates indicated no significant differences in mean effects for programs incorporating simple program designs (3 or less components) than those incorporating complex (4 or more components) designs ($p=.219$).

Instructional strategies were delivered in a diverse number of ways and varied significantly across programs with the most common being forms of tutoring, work-based learning and mentoring. Programs reported significant numbers of support staff beyond the teacher/educator (i.e., psychologists, counsellors, mentors, parents, etc.) as a requirement for program implementation, although there was virtually no discussion of the costs associated with this level of support or the extent to which students did/did not

access those personnel resources. Unfortunately, very little reporting on demographic information within studies reviewed, meant analysis of differences among rural/urban communities, male and female students as well among varying minority and ethnic groups was not possible and a disappointing result. A comparison of mean effect sizes between Canadian and Non-Canadian based programs for physical presences outcomes (dropout rates, absence rates, etc.) showed a statistically larger mean effect for Canadian programs ($g = .247$, $p < .003$), although the small sample size and large variability within studies suggests caution in interpreting those figures.

Almost three quarters of the 151 overall findings indicated evidence of positive program outcomes with another 13% reporting strong positive program impacts. It should be noted however that this finding could be reflective of publication bias, where only positive outcomes are being made publically available/published. Recent discussions in the United States education system underscored this problem with prominent educational associations like the Alliance for Excellent Education arguing that significantly more teens are dropping out of high school than is being reported, making it difficult to accurately assess the impact of resource allocation and educational programming on student performance (e.g., Pinkus, 2006). Correlation analysis revealed a significant and negative relationship between magnitude of effect size (e.g., g) and publication date ($r = -.398$, $df = 64$, $p < .01$) suggesting potentially that research reporting weak or negative results are now more commonly being published and made available, whereas they might have historically not have been. Although there were some indicators of program effectiveness in the studies reviewed, poor reporting and documentation of how instruction was designed and delivered limited our ability to explore statistical

relationships between outcomes and specific program features that might be “best practices”.

Future research should focus on linking outcome impacts (e.g., enrolment status/dropout rates) with studies of contextual change (i.e. wrap around student services, community initiatives, whole school reform, school climate, economic initiatives, etc.) to look beyond “the learner” as the site of dropout prevention. Efforts like the New Approaches New Solutions (NANS) initiative in Quebec which conducts strategic analysis on the impact of school, personal, family, social and institutional factors on school success may in the future provide the kinds of evidence we need to better understand promising practices particularly in disadvantaged areas. Research into how gender, ethnicity, special needs status, SES, and the labour market might moderate effects of instruction is also needed.

Most importantly however, our study highlights the need for Canadian researchers specifically to provide significantly more detailed reporting of the instructional design and delivery of interventions (e.g., specific pedagogical practices, classroom and school climate, individual student characteristics, budget and costs associated with program delivery) and to ensure that rigorous methodological standards are applied to the collection, analysis and reporting of the information. This call was echoed a decade ago when Canadian researchers Anisef and Andres reviewed the School Leavers Survey and argued that “research endeavors... [are] creating a one-sided view of dropouts that [do] not question school policies, practices of schools, the role of the community or the current economic and labor market climate” (p. 97) and that research needs to be placed

within a framework that reports on and includes analysis of individual, family, school and social conditions.

Although this review did not find sufficient evidence to endorse particular strategies and practices to reduce dropout rates, it illuminates several key discussion points for stakeholders and practitioners that reflect the global trends and approaches revealed in the programs reviewed. These recommendations include: a) revisiting existing practices and policies to determine if mainstream classroom practices support the school/work connection; b) examining enrolment and attendance policies and practices for students who have prolonged absences and identifying alternatives for disciplinary action that to not require students to leave school temporarily; c) conducting intentional, systematic reviews on existing programs and policies and making those results available for practitioners locally and in the broader educational community; d) exploring the coordination of services to students placed at risk and procedures for early identification of learners needing extra support and finally e) redefining instructional practices based on best-practices in teaching to accommodate self-direction, flexibility in course delivery and responsiveness to the needs of at-risk learners.

Current Canadian Issues in Dropout Prevention Programming

Canadian Historical Context

The number of students not completing high school is receiving increased attention as a serious challenge facing the North American educational system (e.g., Barton, 2005;

King, 2005). Community, civic and business leaders are concerned about the high percentages of students leaving before graduation without requiring the occupational and academic skills needed to participate in an increasingly high-tech and quickly adapting work environment. Stats Canada reported that during the 2004/2005 academic year 1 in 10 Canadian students dropped out of high school, and while the overall dropout rate in Canada has shown some improvement in the last decade, there is considerable variation in levels of success between provinces as well as between rural and urban districts (King, 2005; Bowlby & McMullen, 2005). While some are quick to point to the effectiveness of dropout prevention programs as the reason for declining dropout rates (e.g., Bowlby & McMullen, 2005; Canadian Council on Learning, 2005) little is actually known about these programs and the kinds of instructional practices they are employing.

“the substantial decline in the dropout rate over the past decade suggests that many of the programs that have been put in place to encourage young people to stay in school until they graduate are meeting with success” – Bowlby, Statistics Canada, 2005

A search of Canadian sources reveals a high volume of locally based attempts to prevent school dropout, increase retention and encourage re-entry of students who have dropped out. The government of Alberta alone identifies several hundred programs as “dropout prevention” programs operating within their province. The proliferation and sheer number of interventions creates a different kind of problem however, creating “confusion and uncertainty as to which group of prevention strategies or which approach to the dropout problem might be most appropriate for any particular school” (MacLean & Janzen, 1994, p. 54). We need to begin systematically examining this body of literature, identifying the kinds of instructional practices and program features that are proving

effective in the context and synthesizing those results in ways that make them accessible to instructional designers, policy makers and educators.

In Canada, as in many other countries, there has been increasing interest in the issue of high school dropout rates. It is seen as a necessary role for schools and related agencies to utilize available resources to reduce dropout rates and ensure that students leaving schools possess the vital skills and experiences they will need to pursue productive and enriched lives. The sections below outline some of the main features of dropout intervention efforts in Canada over the past 16 years, drawing attention to the limits and problems which relate to them.

In the early 1990's the dropout "problem" became a salient public policy issue and the concerns of the Canadian government were serious enough that they launched several reports warning of the negative impact dropping out would have on the individual as well as society as a whole. Nearly two decades later those same concerns continue to be voiced. Frequently discussed are the consequences of not completing high-school and the costs to families and society in not helping students complete basic high school education (e.g., Gordon, 2003; Lecompte & Dworkin, 1991). The credentials assigned to graduates are essentially a form of currency that can be used to gain access to further studies or improved prospects in the labor market but also can be used as a form of social "acceptability" or "conformity".

Well before "global competitiveness" became a public concern in the late 1980's (and perhaps even more pronounced today in educational and political rhetoric) researchers were exploring the economic and social consequences of dropping out. Rumberger (1987) in a review of the literature argued that "by leaving high school prior to

graduation, most dropouts have serious educational deficiencies that severely limit their economic and social well-being throughout their adult lives' (p.101). Over the past forty years there has been considerable analysis conducted on estimating the net benefit of education and has often used a human capital approach; where the costs of education are seen as investment expenditure and earnings/levels of employment are treated as benefits to the investment. In 2000 HRDC reported that employment rates for high school graduates was 10-15 percentage points higher than for dropouts and 20-30 percent higher when compared to those with 8 years of education or less. (HRDC, 2000). In general the research supports the view that dropouts are at a relative disadvantage in the labor market, earning less than graduates, experiencing higher unemployment rates, working in less stable reference jobs, more frequently receiving social assistance (Tanner, Krahn, Hartnagel, 1995). Since the early 1990's researchers have explored the social correlates to unemployment and underemployment and found links to a range of negative effects including alcohol and drug use, crime and delinquency and impaired physical and mental health (Tanner et al. 1995). Dropouts have also been found to experience substantially more personal and family problems (Wagenaar, 1987). However, the causal impact of dropping out on subsequent experiences has not been adequately demonstrated making it difficult to determine whether dropping out or other risk factors which predated dropping out were the main cause.

In response to these concerns the Canadian and provincial governments launched a series of campaigns against early school leaving. Anisef & Anfres (1996) contend that these initiatives grew largely out of a rapidly escalating social belief that dropouts were becoming a problem. This was supported by an increasingly pervasive idea that labor

market development was shifting away from occupations to skills which meant that having high graduation rates would ensure “a skilled and competent workforce” able to “compete in the global market of tomorrow” (Price Waterhouse, 1990, p. i). Interestingly, a decade and a half later, the same arguments are being made, namely “in today’s world, successful completion of high school has become a prerequisite to reaping the benefits of the knowledge economy and a stepping stone for further personal and professional progress; (De Broucker, 2005).

The most prominent initiative in Canada was the five-year \$300 million dollar Stay In School Initiative launched by the federal government in 1990. This project combined a public relations campaign with increasing partnerships between schools, parents, social agencies and youth and the design and delivery of both instructional and non-instructional interventions to combat what were viewed as unacceptable dropout rates. When dropout rates showed little improvement in the few years following, some began to question whether high school non-completion was the result of a myriad of individual influences that are beyond policy intervention and amelioration, or if it was symptomatic of deep structural problems within the school system that could be remedied. Many attributed failure to improve graduation outcomes on the lack of institutional support and the inability of interventions to affect how school at a more fundamental level designed and delivered instruction (e.g., Grannis, 1991).

In the decade that followed provincial governments have also attempted to respond. Reviews of the education system (including how to address dropout concerns) were initiated by the Ontario, New Brunswick and British Columbian governments and the Quebec and Saskatchewan governments have conducted large-scale anti-dropping out

campaigns urging youth to stay in school. Some provinces have also undertaken their own analysis of the provincial-level data from the 1991 national survey of early school leavers (e.g., Alberta Advanced Education and Career Development, 1993).

The concerns of the business community have become an increasingly persuasive voice in the national dialogue on dropout interventions. Critical assessments of the educational and training systems focus on how a poor functioning education system (indicated, among others things, by high dropout rates) is contributing to a perceived decline in Canada' global competitiveness in the labor-market. (DeBroucker, 2005). Reports in the early 1990's (e.g., *Reaching For Success* (Bloom, 1991); *Canada: Meeting the Challenge of Change* (Canadian Labour Market and Productivity Centre, 1993) has echoed this argument. A report published late last year by the Canadian Policy Research Network argued that "Canada has ignored the enormous waste of human potential in these people {early school leavers}, more often blaming the victim than the system of education and the labour market" (De Broucker, 2005, p. i). It appears that the dialogue surrounding dropout issues has remained relatively constant although we have not seen the kinds of performance improvements being argued for.

Despite enormous expenditures in educational funding, the intervention programs and initiatives of the last decade and a half have seen mixed results. Thus current claims in the national dialogue that more work and discussion is needed among those working in the Canadian context so more successful innovations can be implemented (Jordan, 2006). Largely failed attempts from the early 1990's in both the private and public sector may be responsible for the current rise in dropout prevention initiatives. The number of programs explicitly claiming to prevent dropout or help dropouts' transitions back to school are

growing rapidly, both within the public school system and through privately run and non-profit community organizations. Recently for example, the Ontario government launched an \$18 million dollar “Lighthouse Project”, sponsoring 105 new projects (in addition to the myriad of pre-existing projects) aimed at increasing credit accumulation, reducing dropout rates, linking more programs with colleges and encouraging students who left school to return and complete their diploma requirements (Ontario Ministry of Education, 2005).

Unlike many of the initiatives of the early 1990’s there are increasing calls for creating some method of evaluating and distributing information on these burgeoning programs to a broader public. Jordan (2006) argued that “there is no national forum, website, or database where programs that represent best practices, promising innovations and useful ideas can be examined and compared: (Jordan, 2006, p.6). The predilection for talk over action is reflected in much of the broader North American research on dropout intervention which has focused more on counting, describing and classifying dropouts than in evaluating and seeking interventions and solutions. The experiences in the Canadian context since the 1990’s has proven that achieving optimal outcomes within these programs will require “that we learn from past successes and failures and develop new approaches based on a firmer understanding of the etiology of the problem we want to address” (Roderick, 1993, p. 21).

Assessing the state of dropout prevention initiatives that have been undertaken in Canada since 1990 was important within the Canadian context for four main reasons. First, given current pressures to reduce dropout rates, practitioners and policy makers are looking for direction in how to select or design prevention programs that will help them

achieve desired outcomes (e.g., De Broucker, 2005). Maclean and Janzen (1994) noted that a successful dropout prevention program cannot be a haphazard choice made from a series of competing available programs. Rather, educators need guidance and direction to make sense of the existing prevention literature to design specifically targeted and locally relevant instructional practices. More than a decade ago Levine (1992) argued that “secondary schools have been failing for many years to increase the proportion of graduates, yet the program proposals... follow the same models that have already been tried without much success” (p. 268). Years later and we are still looking for guidance and direction on how to design and implement more effective programs

Second, there is increased pressure for educational decisions to be grounded in scientifically based evidence (e.g., NANS/MELS, 2002). Providing funding agencies and organizations with information about how programs operate, how success is being defined and measured, what “benchmarks” for success might look like, and how to best manage the processes and resources for learning is part of reframing the school system around an evidence-based culture that responds to challenges strategically instead of reacting hastily (Smink, 1992). This study aimed to provide a picture of what the research evidence suggests as “best practices” and inform administrators and researchers on how they might best go about collecting data and evaluating their own dropout initiatives.

Third, some researchers are warning that while the overall dropout rates appear to be improving the social and economic implications for those not completing high school is becoming increasingly severe (e.g., De Broucker, 2005). For example, De Broucker noted that labor force surveys conducted with young adults in 25 OECD countries

(Organization for Economic Co-operation and Development) found that education has a much greater positive impact on employment opportunities than working experience. This means that job prospects will not improve much as less educated individuals get older unless they get more education.

Finally, regional and geographical variations necessitate better understanding of how the features of prevention programs may work together to promote school persistence for specific student populations (CCL, 2005). The Canadian context is unique in its efforts to deal with the issue of student retention in that the political realities limit the role of the federal government in shaping educational policy. Provinces are required to inform themselves about “what works” in their own contexts, with their own linguistically, culturally, socio-economically diverse student populations but few provinces support or conduct significant research on the topic (Levin, 1992). Jordan (2006) contended that our efforts to implement strategies to support at-risk learners from dropping out are “confounded by differences among provincial policies and resources” (p.1). This study involved a systematic review of the literature that intended to explore geographical differences in program delivery and design as well as the unique instructional properties of programs designed for specific student populations (e.g., Aboriginal early-school leavers).

Identifying Underlying Factors Relating to Dropout

Students drop out of school for many reasons. Exploring and understanding the factors that go into a student’s decision to dropout is critical in learning how to improve educational programming and support. Reviews of typical characteristics of dropouts, of schools with high dropout rates, and of self-reported reasons for early-school leaving

before graduation reveal the complexity of the dropout problem. As Wagennar (1987) noted;

“the precursors to dropping out, the decision to drop out, the process of dropping out, the responses to dropping out, and the consequences of dropping out all result from a complex interplay of personal, social, situational, structural and contextual factors” (p. 165).

Moreover, if so many factors compound and converge over time, it is virtually impossible to establish a direct causal connection between any one factor and the decision to leave school. Instead, researchers have tended to focus on testing models and exploring the predictive strength of factors that suggest an increased dropout risk (e.g., Goldschmidt & Wang, 1999; Rumberger, 1995).

Researchers have attempted to identify who drops out of school in order to help educators and policymakers develop programs, policies and interventions that promote early identification of students at risk in order to improve graduation rates for those students in particular. Over the years a large body of empirical research has identified numerous factors that predict dropping out and those variables have been well described and comprehensively reviewed in the research literature in other places (e.g., Rosenthal, 1998; Rumberger, 2001; Natriello, 1986; Tanner, Krahn & Hartnagel, 1995).

There are those, however, who strongly oppose the use of predictive modelling, arguing that many who drop out do not fit the profile and many who fit the profile actually complete school on time (e.g., Gleason & Dynarski, 1998). Roderick (1993), using data from a US national survey found that the majority of dropouts had not become so disengaged from school by grade 10 that their withdrawal was inevitable; cautioning that generalizations about a “dropout” profile can be misleading.

Rather than reiterate what previous reviews have detailed at length, only some of the critical distinctions and findings will be discussed here. Two frameworks in particular helped to conceptualize the various correlates and predictive factors discussed in the literature and use that information to inform the review coding scheme. Rumberger (2001) suggests two categories of factors that can be used to explain differences in dropout rates among social groups. First, he suggests that there are a variety of *individual* attributes of the student (e.g., values, attitudes, behaviours, academic engagement, social engagement, past achievement, educational aspirations, self-esteem, gender, race, immigration status and language background) that relate statistically to dropout. Second, emerging from the developmental behavioural science perspective is research documenting that *institutional* (e.g., family, school and community) factors relate to dropping out. Rumberger cites socioeconomic status, strong parental/student relationships, student composition of schools, pupil/teacher ratio, student perceptions of fair discipline policies at school and employment opportunities as some of the key institutional factors that research has demonstrated are correlated to dropout rates.

Lehr et al. (2004) present a different framework, creating a dichotomy between variables which can be influenced to change the trajectory leading to dropout and those which cannot. In their model, *status variables* (e.g., socio economic status, disabilities, family structure), are factors difficult and unlikely to change. However, *alterable variables* (e.g., attendance, attitudes towards learning, identification with school) are easier to change and can be influenced simultaneously by multiple contexts in a students life (e.g., friends, parents, teachers, community members) making interventions more likely to succeed. Recognizing the difference between variables an educator can influence

and those that are relatively fixed features of a student's life is critical when we think about designing and implementing intervention strategies within the educational system. Despite differing models and frameworks for conceptualizing risk factors and correlates, what is clear from the review literature is that dropping out is not so much an event as it is a process often associated with particular individual and contextual variables.

Despite a voluminous research literature exploring and summarizing an extensive list of variables and predictors associated with dropout, none is a reliable predictor of whether an individual student will, at a given time, make a choice to leave school before completion. Dynarski (2001) warns that too often risk factors are used in a simplistic way for determining who should be given priority for prevention interventions without considering how risk factors coexist with factors associated with high academic achievement. Because risk factors in and of themselves may not be strong predictors of dropping out, programs may not be using the best information for determining who to serve (p. 7). The “push and pull effects’ (Jordan, McPartland & Lara, 1999) that characterize the complex decision-making process involved when a student drops out presents challenges for instructional designers and educational programmers, namely whether interventions should address surface issues and attempt to “stem the tide” by increasing attendance in the short term for example, or whether interventions should be designed to address core issues associated with disengagement from school over the long term and work to teach students the skills (academic, behavioural and psychological) necessary to be successful in a school environment. Attempts to address these complex considerations is reflected in the variety of interventions implemented to support students at-risk of dropping out, which we will explore in the following section.

Defining Educational Programming for Dropout Prevention

This research program sought to collect and analyze information on a range of educational interventions that define themselves as a form of “dropout prevention”. Following the general categories of prevention strategies outlined by Morris, Pawlovich, and McCall(1991) and Shannon & Bylmsa, 2005 how models of “dropout prevention” are operationalized in the literature was reviewed. Models included: a) strategies that attempt to prevent the development of factors which correlate with early school leaving (e.g., pre-school early intervention programs); (b) strategies designed as interventions that support students “at-risk” of dropping out (e.g., alternative instructional programs, mentoring, curriculum change to increase relevance); (c) strategies for school-to-work transition that attempt to persuade students to stay in school by building connections between the learner and the labor market (e.g., cooperative education programs, work orientation workshops, career awareness programs); and (d) re-entry strategies that attempt to retrieve students who have left school by offering specialized instructional and support programs (e.g., alternative schools, literacy programs).

EARLY PREVENTION PROGRAMS

Many researchers contend that at-risk students arrive to their first classroom experience in school already demonstrating or experiencing learning deficits or social difficulties that put them at risk for early school leaving in later life. Early prevention programs are designed to target student populations who share characteristics that research has shown correlate with early school leaving (e.g., low socio-economic backgrounds, speak a primary language other than the primary language in the school (usually English or French), are part of an ethnic minority) (Shannon & Bylmsa, 2005).

Thus, pre-school and early childhood prevention programs have been designed to intervene early enough to provide these children with the skills needed to achieve academic success. Schargel & Smink (2001) argue that providing the best possible classroom instruction from the beginning is the most effective way to reduce the number of students who may drop out of school (p. 41).

Conclusions as to the effectiveness of these programs have varied across studies and across time but recent investigations have increasingly been reporting positive effects for children at-risk. For example, in a study conducted by Rumberger, (2001) one-third as many students who attended the High/Scope Perry Pre-School program as non-preschool program members graduated from high school or received their GD (71 percent versus 54 percent) (p. 26). He concluded that because these outcomes were measured more than a decade after the intervention ended that “early interventions for persons at-risk of dropping out can be effective” (p. 26). These strategies can take many forms, including: elementary school remediation or resource support programs, elementary school guidance programs, parenting programs, home-school liaison initiatives, parent involvement programs and other pre-school programs.

INTERVENTION PROGRAMS (JUNIOR AND SENIOR)

Supplemental support programs make up the majority of programs aimed at addressing the issue of early-school leaving. These programs are created to support students who have been identified as being “at-risk” of dropping out and generally fall into one of two categories: academic support programs and out-of school enhancement programs.

Academic support programs tend to focus on providing students with more of what they are already engaging in; more homework support, more instructional time, double class periods, before and after school homework clubs, Saturday school or summer school. The purpose of these programs is to help student successfully complete the core courses required for grade progression and to help them manage the content being taught. School-based programs can also focus on whole-school changes that they believe will be particularly beneficial to student at-risk. Curriculum changes to increase relevance and encourage student engagement and social-bonding as well as attendance monitoring programs are a few examples of these kinds of initiatives. These programs can also involve incentives for attendance (e.g., Epstein and Sheldon, 2002) or include access to additional resources like technology, individualized instruction plans, field trips, etc. Morris et al., 1991 listed at least 14 different interventions that researchers in the early 1990's had found evidence indicating a positive impact on at-risk students.

Out-of School enhancement programs are designed to support the variety of unmet needs students have outside of academics which are believed to play a role in a student's decision to leave school. Programs are frequently implemented to address students' mental, physical and social needs, hoping that by providing additional resources and sustaining these services over time to at-risk students they will help increase retention rates and improve student performance. Service learning, mentoring, tutoring programs, peer-counselling, adventure training, drug and alcohol addiction support programs; and extra support services for teenage parents are all examples of out-of school initiatives. Research evidence demonstrating the success of these programs is mixed. Dynarski (2000), in a study comparing students in a supplemental pull-out program to a control

group found that “supplemental programs had almost no impacts on student outcomes....dropout rate, and average student grades. . . were similar among treatment and control groups”(p. 3). Others however have suggested that these programs can have lasting positive effect. The “Change Your Future” program run in the Toronto Board of Education in the early 1990’s focused on individual counselling and group meetings as an intervention strategy for middle school students at-risk of dropping out. The researchers found that dropout rates for students in the program were lower then students in the nonprogram group as were transfer rates and credit accumulation (Brown, 1994).

SCHOOL-TO-WORK TRANSITION STRATEGIES

Since one of the main concerns over lingering dropout rates is the economic impact that lack of education will have on an individual and ultimately on the workforce in general, it is not surprising that partnerships between schools, communities and industry have been producing a wide range of programs designed to help students stay in school and prepare them for the transition to the workforce. These “transition” programs tend to focus on career preparation, job counselling, co-operative education and career information centers. In the early 1990’s Bloom produced a review of 30 “successful” business-education partnerships in Canada that explicitly intended to enhance student retention in high schools. He found very diverse approaches across the country including cooperative education, job shadowing, apprenticeship programs, alternative programs for student working-part time, interactive videos on career options and computerized career information systems among others. All of these Canadian programs were reporting some measure of success in enhancing student retention (Bloom, 1991).

RE-ENTRY/RETRIEVAL INITIATIVES

The final category of programs addressing the issue of dropouts are those that seek to reconnect with students who have already dropped-out by offering them a “second-chance” to complete their formal schooling. A common type of retrieval program is the establishment of an alternative school which aims to address the needs of students, who for whatever reason, seem unable to have their needs met in the regular school system. Schargel and Smink (2001) argued that alternative schools tend to share some common characteristics including; (a) low teacher-student ratio; (b) small number of students in the school population; (c) individualized learning programs, (d) administrative and school staff that set high expectations for student achievement, and (e) a flexible school schedule with community support. (p. 117). Recently cyber schools and virtual schools have been emerging as an additional option for students considering re-entry into the school system. Research on the effectiveness of these programs and the extent to which issues of accountability and quality of education are being addressed remains unknown (Shannon & Bylmsa, 2005).

Characteristics of Effective Prevention Programs

Much of the debate on “best practices” for dropout prevention programming has tended to remain in the realm of abstraction and speculation. While detailed lists and “key features” of effective programs are readily available (e.g., National Dropout Prevention Center: <http://www.ed.gov/programs/dropout/dropoutprogram.html>) to those working to design and manage programs, assertions about the value of programs and the characteristics that make them effective have traditionally been “based on a few or several descriptive statistics...we do not always find empirical evidence to back the

claims being made” (Morris, Pawlovich,& McCall, 1991, p. 67). The growing body of anecdotal evidence and case studies has contributed to the general consensus among researchers as to several key and promising practices for dropout prevention programs. Researchers caution however that given the vast array of program designs its clear there is no single right way to intervene in all cases. The table below lists the key successful components listed by several highly referenced sources and reveals consistent overlap in the types of strategies being defined as “effective”.

Table 1. Key Characteristics of Dropout Prevention/School Completion Programs

Author	Main Findings
Dynarski (2001) (2000)	<p>Creating small schools with smaller class sizes</p> <p>Allowing teachers to know students better (relationship building, enhanced communication)</p> <p>Provision of individual assistance</p> <p>Focus on helping students address non-academic needs through access to family/personal support services</p> <p>Oriented toward moving students to obtain GED certificates</p>
Rumberger (2001)	<p>Create a non-threatening environment for learning</p> <p>Are staffed by caring and committed members who set high expectations for student success and take personal accountability for student success</p> <p>Create a culture that encourages risk-taking, self-governance and professional collegiality, and</p> <p>Provides a flexible school structure with a low student-teacher ratio to promote student engagement</p>
West (2001)	<p>Small student population to allow for meaningful relationships</p> <p>Extending the role of “teacher” to incorporate mentoring and counselling</p> <p>An explicit schedule to support effective governance</p> <p>Clear linking between school and what is used in the workforce</p> <p>Instructional practices that are responsive to student interests/needs</p> <p>Close monitoring of student behaviour and academic progress</p> <p>Recognition and reward for improving performance</p> <p>Activities that contribute to students’ sense of worth and value</p> <p>Administrative practices that encourage teacher collaboration, innovation and accountability</p>
Schargel & Smink (2001)	<p>Early intervention (i.e., early literacy, family involvement)</p> <p>Promoting opportunities for the students to form bonding relationships (i.e., mentors, service learning, etc.)</p> <p>Providing opportunities for professional development, diverse approaches to learning styles, using technology and individualized learning</p> <p>Making the most of the community through community collaboration, career education, conflict resolution and programs to enhance inter-personal skills</p>

1. Dynarski (2001) study funded by the School Dropout Demonstration Assistance Program
2. Dyanarski (2000) A summary of evaluations of alternative programs funded through the Federal Dropout Demonstration project
3. Rumberger (2001) A review of literature on why students drop out and the types of strategies implemented to reduce dropout rates.
4. West (2001)A review of the literature
5. Schargel & Smink (2001) Review based on program database generated by t he National Dropout Prevention Centre

Taking advantage of the experience and knowledge of others is an important strategy in the creation of new educational ventures like dropout prevention program as well as the modification of existing interventions. How this is done varies considerably from program to program. Program planners may carry out scholarly and systematic research while others may make decisions based on personal experience with comparable programs or on reports from outside consultants. Lehr et al., (2004) contends that rather than searching for the perfect program, “identification of components that facilitate the effectiveness of interventions may prove a more valuable endeavor” (p. 18). By examining a range of interventions across a variety of contexts this project sought to identify those key components to help guide the development of interventions, improve the likelihood of successful delivery and provide a framework for evaluating outcomes in the future.

There have been a number of published reviews on dropout prevention programs in the last 20 years. However the evidence of their effectiveness has been weak for several reasons.

First, many reviews have focused on highly specific target populations (e.g., young mothers, school-aged children, young offenders, students with disabilities, etc.) making the generalizability of findings to the broader population difficult to ascertain (e.g., Somers & Piliawsky, 2004; Reyhner, 1992; Cobb et al., , 2005). This study proposed an inclusive approach to selecting studies for review and then coding the specific features of the programs to allow for comparisons across the range of programs and within programs targeting specific populations.

Second, researchers caution that while there are many reports describing promising dropout interventions, only a small portion of these offer conclusions that are based on rigorous empirical research (Steinberg & Almeida, 2004; Shannon & Blymsa, 2005; Dynarski & Gleason, 2002). Cobb et al., (2005) further contend that “most prior reviews [of prevention programs] are exclusively narrative reviews with no attempt to screen studies with weak designs” (p. 10). While the varying quality of primary studies cannot be argued, this review coded for and reported on the methodological make-up and quality of the studies being produced and reported. This assisted in (a) providing explicit direction and advice for future researchers and program evaluators on how to improve the methodological rigor of their work as they collect and report on research findings; and (b) will provide us with a way of examining how different methodologies and approaches might result in different kinds of interpretations about how dropout-prevention programs operate.

Third, prevention programs that have been or are currently operating in the Canadian context are virtually non-existent in the current review literature. Jordan (2006) reported that to date in Canada there is “no national forum, website or data base where programs that represent best practices, promising innovations and useful ideas can be examined and compared...more are needed” (p. 6). This study provided a comprehensive accounting of the publically available research evidence on prevention initiatives in Canada since 1990 and explored how the delivery of those programs was undertaken in our own national and provincial contexts.

Finally, previous reviews have been concerned mainly with the question “do they work?” and have not synthesized the findings on the more instructionally relevant

question of “how do they/do they not work?” Several researchers have recently pointed to the need for systematic reviews on this topic to take into account the variety of instructional features and contextual realities (e.g., community economics, community support, value orientations) that may moderate program effectiveness (Aron & Zweig, 2003; Steinberg & Almeida, 2004). Rather than just synthesizing findings, this study involved coding of studies to explore the larger context in which programs operate and the variety of instructional and non-instructional features that may hinder/promote its success. A good review can also address the “how it works” concern.

As one researcher noted: “we do not yet have a menu of program options for helping students at risk of dropping out. The evaluation findings are useful as guides ...but they fall short of providing a scientific basis for implementing programs in new schools or districts based on the models” (Dynarski, 2004, p. 265) What Works Clearinghouse, an organization which applies rigorous review standards in identifying effective programs and practices, listed dropout prevention as one of their research priorities in 2002 and a recent report to the US Congress senior researchers contended that “providing funds to support more rigorous evaluations of existing programs would help identify the most promising programs and practices for preventing dropout” (Shannon & Blymsa, 2005, p. 60). Some researchers have suggested that the smaller number of empirical research reports produced in Canada in the last two decades makes it difficult to examine and compare programs and provide direction for educational policy (Jordan, 2006). Many argue however that despite the challenges involved in working with this methodologically diverse body of literature, “much can be learned from the research studies and reports that can be applied to schools and classrooms” (Shannon & Blymsa,

2005, p.11). There was a clear mandate to begin systematically investigating and synthesizing this research literature.

The Role of HPT in Public Education Interventions

Human performance technology (HPT) is a relatively recent convergence of a broad range of theoretical and applied disciplines such as economic theory, cognitive science, systems theory, learning theories and management science, among others (Combs & Falletta, 2000). Though several working definitions exist on what human performance technology means most emphasize the field as a set of methods, procedures, and strategies for solving performance problems or realizing opportunities related to the performance of individuals, small groups and large organizations. Practitioners in the field seek to realize performance opportunities in military, industry, and educational systems through the design, selection, implementation and evaluation of appropriate interventions (Surgue and Fuller, 1999). This project focused broadly on processes associated with identifying explanations for program successes and disappointments (i.e., cause analysis) as well as testing assumptions about effective intervention factors (i.e., intervention selection).

Systems and organizations continually experience change as adaptations are made in response to local pressures and short-term performance goals (e.g., cost, fluctuating dropout rates). People adapt to their environment or they change their environment to better suit their purposes. Several decision makers striving locally to optimize

performance at different times in different parts of a program's administrative structure often creates opportunities for performance problems to occur. Cause analysis is a structured step-by-step process that helps identify what, how and why performance problems occur, so they can be solved and avoided in the future. At its core cause analysis is interested in questions relating to how performance problem occur, what environmental factors contributed to creating it, as well as when and where problems occurred (Mager & Pipe 1997). In the context of this project, the goal was to broadly explore the mediating and moderating variables that may impact on a dropout prevention programs ability to meet desired performance goals to explore questions of how we might adjust practices to reduce the impact of those variables or design supports to limit the impact those variables have.

Intervention selection involves a systematic, comprehensive, and integrated response to performance problems and their causes as well as to performance improvement opportunities. Administrators commonly rely on what Poister (1978) referred to as “back of the envelope” analysis whereby a thumbnail analysis of the issues or quick evaluations are made based on judgments or intuition combined with accumulated experience. Although these assessments are sometimes very accurate and often are the only expedient course to take, such efforts often lead to program delivery based on facts whose validity is untested. Intervention selection seeks to answer questions related to which interventions, under which circumstances would be the most appropriate in terms of cost, available resources, organizational culture etc. This project intended to create a starting point for testing assumptions about dropout program effectiveness and developing more complete information on which designers and administrators can base

decisions related to which kinds of interventions may work in which contexts. For example, understanding how employees (e.g., support staff), resources (e.g., type of curriculum) and the organizational environment (e.g., philosophical approach to the program delivery) factors for dropout prevention programs with documented successes were organized, helps provide information on where resources could be allocated in a certain context or what kinds of program feature (or combination of program features) seemed to produce the most positive performance changes.

Politics and Purpose: HPT in Traditional Educational Contexts

Although human performance technology (HPT) has more traditionally been associated with corporate training within private organizations, the application of HPT principles to publicly run educational systems and programs dates back to the late 1970's when researchers were exploring how, for example, an Instructional Quality Profile tool for training might apply to investigations of curricular interventions in schools (Merrill, Reigluth & Faust, 1979). More recently, a highly publicized book, *The Eden Conspiracy: Educating for Accomplished Citizenship* (1998) written by Dr. Joe Harless, a renowned performance technologist, advocated for the use of human performance approaches in creating an accomplishment-based curriculum for school improvement.

Although some may contend that public and private corporations experience the same kinds of obstacles when moving toward implementing models and theories of human performance improvement, “the public/political sector is quite different, so the logic and rationality that may apply to a private-sector body cannot easily be extrapolated to them” (LaPalombar, 2001, p.558). LaPalombar takes a critical view of trying to ‘copy’ what the private sector does saying:

"These theoretical frameworks may work quite well for the private sector, where one finds much clearer statements of purpose or of means and ends and where the boundaries demarcating organizations, their authority, their responsibility are much more unambiguously delineated" (p. 558).

The paradox here is that on the one hand we want effective dropout prevention programs where people work together to achieve common learning goals, where collaboration, sharing of information, creative and innovative work is carried out; but on the other hand, these programs are nested in a larger community which places competing demands on them. Like many other public institutions schools and educational programs in general are expected to produce or manage very vague, diffuse contradictory and even conflicting performance priorities (Levin & Sanger, 1994). Schools are "collectivities whose participants are pursuing multiple interests, both disparate and common, but recognize the value of perpetuating the organization as an important resource" (Scott, 1998, p. 26). Such is particularly the case when dealing with programs and schools that seek to address the myriad of issues related to students at risk of dropping out. It is the pursuit of multiple interests that makes the application of human performance improvement models often difficult to transfer from business to school-settings. To say that private-sector organizations pursue only one goal, the bottom line, would be an oversimplification, however one thing is clear, the purposes of schooling are much more diverse, culturally relative, dependent on the neighbourhoods, circumstances, student populations, personnel and politics (Collinson et al., 2001) than are private sector organizations.

To date, little has been done to bridge the gap between approaches and techniques associated with human performance technology and the kinds of educational dilemmas and performance problems being experienced and observed in traditional educational environments dealing with students at risk of dropping out. However, public administrators are demonstrating growing interest in HPT as a means of qualitatively improving the services and educational programs that they run. HPT today is considered particularly useful for educational administrators and policymakers, as an effective framework from which to analyze and improve the quality of educational programming (Harless, 1998).

Accountability vs. Autonomy in Designing Educational Programming

Another area of consideration when dealing with dropout prevention programs as a human performance issue is how we are to address the contradiction within the educational sector between a desire for creative, innovative and autonomous educational programming and our need for some form of accountability, structure and leadership. Just as in the private sector where “employees search for boundaries of organizational expectations within which to exercise their creative potential, and feel lost and chaotic when such structures are not in place” (Laiken, 2002, p. 6) schools and educational programs are caught in the same dilemma. Ideally dropout prevention programs would engage all related individuals (students, parents, teachers, and administrators) in meaningful, experiential and groundbreaking instructional experiences but there is also increased pressure for programs to demonstrate their ability to reduce the bottom-line dropout rate and improve student performance in the short-term.

In the private sector we often observe a culture in which the “the leader/manager... [helps] employees to learn through calculated risks and careful experimentation. A climate of 'no blame' allows workers to make provisional attempts, receive feedback from supervisors and colleagues, make changes and try again” (Laiken, 2001, p.7).

Traditionally models of dropout prevention programs rarely describe the kind of risk-taking culture evidenced in private sector organizations and programs. As is typical of many forms of educational evaluation that seeks to research how a particular program might improve student outcomes, people often react - no body wants to have their new program serve as an “example”, for good or bad. We can see from recent trends in the US and in Canada that there is growing interest in making educational programming and instructional interventions more "accountable" for student progress. As we work to make connections between traditional learning environments and principles and practices associated with HPT we need to explore how successful programs are designed and accountability structures and innovation are managed.

Program Design and Delivery: Fidelity to Implementation Plans

People would generally agree that the success of educational programming is determined in large part on how stakeholders design the intervention (Lawler, 1996). Thus the link between how we design programs and how individuals within those programs perform cannot be understated. Interestingly, however, there has been little effort to illuminate some of the issues related to program design and delivery as they relate to dropout prevention initiatives. The dilemma we face in addressing the effectiveness of dropout prevention is that we design programs to be successful in reducing dropouts but in the real constraints placed on resources, support services,

instructional time and student access, truly innovative approaches to dropout prevention is a misnomer.

“Just as “an organization that is designed to do something well for the millionth time is not good doing something for the first time” (Galbraith, 1982, p. 98), dropout prevention programs designed to intervene in the same way to all students, in the same physical set-up, using the same kinds of assessment strategies, using the same resources, year after year are probably not the best candidates for demonstrating “best- practices”. What this basically comes to is that we frequently haven't designed, at the most basic levels, dropout prevention programs to be proactive in assessing gaps between their actual and desired outcomes and in fostering a climate where success and challenges are shared and discussed and a culture on continual performance improvement is valued.

Literature in the field of human performance technology provides a very expanded view of how we might look at program design; including things like to (a) external reward systems (Galbraith, 1982); (b) role descriptions and layers of authority (Neil & Wykowski, 1999); (c) the separation between program operations and innovators within the organization (Galbraith, 1982); (d) ‘safe zones’ for research and development (Galbraith, 1982); (e) flexible job boundaries (Sugarman, 1997); (f) cross-trained employees (Sugarman, 1997); and (g) employee salary. Yet, none of the research conducted in the educational sector has explored the contradiction between how dropout prevention programs are designed to function and what we are asking individuals within these programs to produce, how we are asking them to perform and the conditions under which they are expected to perform. Informing investigations of dropout prevention programs with fundamental principles related to HPT requires us to consider how the

system within which these programs operate, conceives of and assigns responsibilities to different stakeholders. How do the unique program features, reward systems, resource allocation patterns, etc. impact on these programs' ability to reduce overall dropout rates? How do the non-instructional features of the programs (e.g., level of non-teacher support staffing, funding source) relate to evidence of program effectiveness? This project seeks to explore these kinds of questions by reviewing the broad range of contemporary program designs and delivery strategies and examining their relationship to program success.

HPT and the Use of Systematic Reviews

Evaluation research methods including systematic reviews, provide a means to ascertain the worth or value of a performance improvement initiative, to improve a performance-improvement processes or to decide to discontinue the effort. It can also be useful in judging the relative worth of performance-improvement alternatives. Pershing (2006) presented five ways in which evaluation can be used in HPT, the final one being as a means of research. He argued that research is:

"the most important and often the most overlooked role that evaluation can play in the HPT process. As performance-improvement initiatives are evaluated, data can be collected and analyzed in ways to add knowledge to HPT principles and practices and provide knowledge that can be generalized to future efforts. Evaluations can be designed to provide reliable and valid data to ascertain what worked, when, and how it worked. This information, when shared with the field, will advance practice" (p. 26).

In the case of synthesizing educational research on dropout prevention programs the HPT perspective complements and strengthens typical approaches to educational evaluation and systematic review in several ways:

HPT emphasizes exploring performance problems within the context of a system, recognizing the interdependency of the various factors that affect performance. It promotes understanding and analyzing these factors before making recommendations on effective and efficient interventions that should be employed (Rothwell & Kazanas, 2004). Reigluth (1999) argued that an important direction for the field of HPT is working to provide prescriptions for more macro organizational methods, particularly in structuring and sequencing instruction as well as the importance of synthesizing these ideas. This project aims to use both of the strategies outlined by Reigluth; analyze (i.e., break down instructional methods into components to determine which ones make a difference) and synthesize (i.e., identify the best combination of instructional methods for a given situation) to develop understandings of “best practices” not only in terms of *what* these programs should offer but *how* best to structure and organize those components within different contexts to achieve desired performance outcomes. Recent advancements in synthesis methodology highlight the importance of understanding how changes in performance are contingent not only on the generative mechanisms of programs and interventions but the underlying resources and contextual variables. By coding and analyzing studies across multiple measures, this project will seek to increase validity by obtaining a more complete picture of how cases compare on a spectrum of measures (Slavin, 1986). The project supports the notion that research synthesis should seek to extract information on what works for whom in what circumstances.

HPT seeks to achieve optimal human performance through strategies and interventions that have been derived from scientific research (Marrelli, 2005). More than a decade ago Clark (1989) was suggesting that one of the pitfalls of instructional

technology is that reviews of prior research are often poorly conceptualized. He contended that research in instructional technology needed to better support building “an accurate picture of the ‘state of the question’ ... before hypotheses are generated or treatments are explored. A researcher must begin with the problem to be solved and uncover the best and most current state of knowledge about possible solutions” (p. 59). Recently, researchers and practitioners in the field have continued to echo these concerns, suggesting that as the field of HPT advances care must be taken not to downplay the effort and time required to analyze performance problems before prescribing solutions. Stolovitch (2000) argued that “practice without solid research and theory base becomes little more than ‘craft’” (p. 13). The nature of systematic reviews is that they seek to select and critically appraise relevant research, to collect and analyze and summarize the data from the studies that are included in the review in order to inform practice and advance theory. Oliver, Harden, Rees, Shepherd, Brunton, Garcia & Oakley (2005) contend that “systematic reviews have become an important tool for facilitating evidence-based policy and practice as they bring together and combine findings from multiple studies” (p. 429).

HPT defines human performance in terms of results not activity. It generally does not support enthusiastic, unsubstantiated interventions that cannot demonstrate firm theoretical foundations or valid performance results (Rothwell, Lindholm & Wallick, 2003). While recent discussions on research synthesis highlight some of the difficulties in combining different types of research evidence which fundamentally differ in their views of what represents “valid” evidence (Oliver et. al., 2005), there is general agreement that systematic reviews which seek to inform public policy and act as an input into decision-

making processes should not be based on studies which report on intuition, impressions, casual observation or “conventional” wisdom but rather studies which offer some reasonably sound evidence for reported outcomes (Cooper & Hedges, 1994). As “evidence-based practice” becomes increasingly cited as a goal of school boards and educational organizations involved in developing and administering dropout prevention programs it is critical to both provide accurate information on the effectiveness of current programs and provide instruction on how programs in the future might better make use of their own performance results to further achieve desired outcomes. The development of systematic measures of the impact of programs is often neglected by educators, not surprising given the political and administrative challenges that accompany evaluative practices. Working to support evidence based approaches to dropout prevention is critical not only for managing change but “validating teaching strategies that often appear iconoclastic and questionable to decision-makers and members of the public unfamiliar” (West, 1991, p. 55) with how these programs operate.

The application of research synthesis in performance technology has some clear complementary connections both theoretically and in terms of ultimate objectives in improving performance and achieving desired outcomes. Marrelli (2005) contends that research reviews are an important tool for the performance technologist, particularly in identifying and developing interventions (i.e., dropout prevention and intervention programs). She suggests that these kinds of analysis “can provide ideas for interventions as well as guidance in the factors to consider in evaluating the potential effectiveness of an intervention in a specific context” (p.42). Throughout the proposed study, the principles of HPT described above informed and directed methods for data extraction and

how results might best be reported to support future instructional and program delivery/design

Purpose & Research Questions

The last major review of dropout prevention initiatives targeted for the Canadian audience was published in 1991 by Morris, Pawlovich, and McCall under the mandate of the Canadian Education Association but it based its recommendations almost exclusively on American studies and did not look to explore potentially unique findings within the Canadian context.

Contemporary researchers note that significant differences in the delivery and oversight of American and Canadian educational systems (for example education in Canada is under the complete jurisdiction of the territories and provinces) is a major reason for encouraging us to look more locally at “what works” because we often do not have the same values, structures and funding channels to support educational innovations (Robertson, 2004, Adams 2003). For example, a large number of the dropout prevention initiatives in the United States are funded through special federal funding made available through the *No Child Left Behind*- Title 1 Grants. It should be argued that part of exploring “what works” is also examining whether those programs require significant infusions of additional financial and human resources and evaluating the likelihood that such practices would be implemented even if they were identified as “effective” (Morris, Pawlovich & McCall, 1991). Given the differences in educational governance and policy making, exploring “what works” in the Canadian context was intended to add to a better understanding of our own unique (perhaps, or perhaps not) circumstances and challenges.

Further, while the USA has a highly organized network of research centres established to explore and research educational programming specifically for dropout prevention (e.g., NDPC/N- National Dropout Prevention Center/Network) as well as the federally funded and operated What Works Clearinghouse which is mandated to help program developers with designing and carrying out rigorous educational evaluations, Canada does not have the same level of oversight on the issue of early school leaving. Greater collaboration and awareness of what individual jurisdictions in Canada are doing to instructionally support potential dropouts is badly needed (Jordan, 2006) and this research attempts to bring together that unexplored literature in Canada.

Although the issue of the “dropout crisis” has been used within the Canadian educational discourse for over 40 years, federal recognition and funding of large scale response to the issue began in the early 1990’s at around the same time that the Morris, Pawlovich, and McCall (1991) review was published. Thus, the logical timeframe to begin including studies for this project would be 1990. Oftentimes, there is a delay between time an article is published and the time it appears in an electronic database. To align hand searches and electronic searches Dec. 2006 was selected as the end point for study consideration.

Thus this systematic review of research on dropout prevention programs in Canada between 1990 and 2006 was conducted with the following purposes in mind:

- Increase awareness and knowledge of current research on educational programming for dropout prevention
- Increase awareness and knowledge of interventions that show evidence of effectiveness in improving student performance

- Increase awareness and knowledge of instructional and non-instructional conditions that may moderate performance improvement
- Strengthen communication between stakeholders through development of common understandings of “success”; and
- Improve decision-makers’ ability to strategically plan for and design programs for dropout prevention, make decisions related to budgeting and staffing and manage available resources that take into consideration current research evidence.

More generally I sought to answer the following questions:

- Has dropout prevention programming in Canada resulted in an increase in student retention rates and improved student performance? If so, to what extent?
- What program and instructional conditions moderate the effects on dropout rates and other performance outcomes?
- What are some optimal instructional conditions or “best practices” for effective dropout programming?

Methodology

Systematic Reviews

As the number of studies being generated in this topic increases, researchers are now interested in exploring new and expanded questions that require looking at these pools of data in ways beyond their original intent (Thorne, 1994). The challenge now is to make meaning from the vast array of information on dropout prevention programs that is now available. A review of the literature on dropout prevention programs for youth at risk in Canada reveals considerable variation in program design and implementation. Moreover, examinations of program effectiveness have yielded inconsistent results. Given the nature

and instructional diversity of these programs and the varied populations they assist, it is unlikely that questions relating to the overall impact of these programs could be answered by conducting another evaluation study of an individual program. Research synthesis provides a more realistic strategy for improving the design and delivery of programs in the future.

As with most methodologies related to secondary analysis of primary studies it is important to remember that we are dealing with a continuum of definitions and approaches. Different researchers have very different definitions and approaches to understanding what constitutes a “systematic review”. A first important point of discussion is about whether these reviews are interpretive analyses or aggregative analyses. Some researchers contend that systematic reviews are valuable because of their ability to build theory through the aggregation of findings. This approach is characteristic of the efforts of the international Cochrane Collaboration who strongly emphasize the use of data from randomized controlled trials in their protocols for systematic review. They contend that just as data aggregation is appropriate at the level of a single study so too is aggregation of findings from multiple studies. Others contend that systematic reviews should focus more on comparing and interpreting than on simply cumulating aggregate findings (Dixon-Woods, Bonas, Booth, Jones, Miller, Sutton, Shaw, Smith & Young, 2006). Given the array of methodologies used for secondary analysis, examples of which are outlined in Table 3 it is clear that defining what constitutes a “systematic review” is complex. There are however common features of most definitions that were used as a frame for this study. Systematic reviews tend to:

(a) be driven by well-focused and feasible questions

- (b) employ explicit procedures or review protocols and methods for evaluating source material
- (c) provide transparent descriptions of methods used so that at least in theory another researcher could reproduce the study and arrive at the same conclusions
- (d) operate as efficient information management tools by providing a way of reducing the volume of information on a topic; and
- (e) are concerned with having practical value to the research community and other stakeholders

Table 2. Types of Secondary Research Reviews

Methodology	Definition	Example
Narrative Review of Research	A narrative report of an intuitive aggregate of individual research findings. These reviews normally ignore unpublished research.	Cook, Mulrow, Haynes, 1997
Narrative Explanatory Synthesis	Narrative reviews deal with a broad range of issues related to a given topic rather than addressing a particular issue in depth. Narrative reviews are often employed if research is scant or preliminary or if studies are very limited by flawed design or execution.	Light and Pillemar, 1984
Meta- Evaluation	The process of delineating, obtaining, and applying descriptive information and judgmental information - about the utility, feasibility, propriety, and accuracy of an evaluation and its systematic nature, competent conduct, integrity/honesty, respectfulness, and social responsibility - to guide the evaluation and/or report its strengths and weaknesses.	Stufflebeam, 2001
Meta-Analysis	The statistical analysis of a large collection of analysis results from individual studies for the purpose of integrating the findings	Glass, 1976
Exploratory Literature Review	This involves scanning the breadth of literature available on a given topic without having specific codes predetermined or without having specific contexts, populations, and methodologies in mind. This is often a very preliminary for so synthesis.	Pascarella & Terenzini, 1991

Grounded Meta-Analysis	A process whereby (a) an open-ended coding instrument is developed and revised as synthesis proceeds; (b) coding is done of both quantitative and qualitative information to permit comparisons across studies; (c) the process of gathering studies, coding information and analyzing the data is recursive and comprehensive; (d) each study is evaluated in terms of quality and low quality studies are excluded; (e) use of debriefers to assure consistency in analysis.	Hossler & Scalese-Love, 1989
Meta-Synthesis	Systematic review in which primary studies are identified using rigorous search strategies, and the focus of the analysis can be either to interpret and describe or to build theory. This is in contrast to hypothetical-deductive reviews.	Bair, 1999
Synthesis of Qualitative Findings	Synthesis of findings across different studies conducted by different investigators. In these studies the authors produce narrative or theoretical combinations of studies in the same topical area, using techniques such as reciprocal translation of key metaphors.	Jensen & Allen 1994
Best-Evidence Synthesis	Incorporates the statistical rigor of meta-analysis to synthesize quantitative findings together with the flexibility of traditional reviews. Statistical analysis is supplemented with rich review of the literature that describes and interprets discrepancies and summarizes results that cannot be quantified.	Slavin, 1986
Systematic Review	This review comprises (a) a literature search for primary studies with comparable aims, methods and outcomes; (b) the evaluation of relevance and quality of each study using predefined criteria; (c) abstraction of data on methods and results	Boaz, Ashby & Young, 2002
Case Survey Method	Researchers use a conceptual framework to construct a set of highly structured questions to collect information from individual case studies in specified topical domains. Answers to these questions are then turned into data amenable to statistical analysis	Lucas, 1974
Qualitative Comparative Method	Boolean algebra is the basis for creating categorical information on key variables across individual cases. A holistic view of individual cases- as distinctive configurations of association, causes and outcomes- is maintained.	Ragin, 1987

Qualitative Meta-Synthesis	Cross-case generalizations created from generalizations made from and about individual cases.	Schofeild (1990)
Meta-Ethnography	Enables a rigorous procedure for deriving substantive interpretations about any set of ethnographic or interpretive studies It compares and analyzes texts, creating new interpretations in the process.	Noblit & Hare (1988)

Conducting a systematic review based on the principles of research synthesis, defined by Cooper and Hedges (1994), namely; “an attempt to integrate empirical research for the purpose of creating generalization” (p. 5) allowed for critical analysis of research on dropout prevention as well as exploring implications for designing instructional and non-instructional program components to best support program success. This approach is similar to that outlined by Light and Pillemer (1984) who advocated for an “alliance of evidence” (p. 50) by incorporating both the quantitative and qualitative data from primary research studies to explore mediating and moderating variables that may impact on how results are interpreted. The EPPI centre outlines five major stages that researchers go through in conducting a systematic review which were used to guide this proposed project. These stages are outlined in Table 4 and described in more detail in the sections below.

Table 3. Phases in Conducting a Systematic Review (EPPI guidelines)

Phase	Procedures
<i>Phase 1.</i> Planning the review	User Involvement Setting up Support for review Setting the scope of review
<i>Phase 2.</i> Gathering and Describing Research	Searching for studies and managing reports Applying inclusion/exclusion criteria Keywording and descriptive mapping Refining the scope of the review
<i>Phase 3.</i> Analyzing and Synthesizing the Data	Extracting Data Quality Assessment Of Evidence

Synthesizing The Findings And Quality	
<i>Phase 4.</i> Writing the report	Drawing up recommendations Developing final report
Disseminating the report	
<i>Phase 5.</i> Making Use of the Report	Reflection and contribution to methodology

The Value of Systematic Review for Dropout Interventions

There are several ways that multi-method systematic reviews help us address issues within complex research areas such as dropout prevention programs in Canada:

MAKING THE UNMANAGEABLE MANAGEABLE

The need for re-examining and disseminating information on available studies is underscored by recent comments expressed at the Canadian conference on creating a National Dialogue For Students At-Risk in early 2006 and recently published by Jordan (2006). She argued that because there is “no national forum, website or data base where programs that represent best-practices, promising innovations and useful ideas can be examined and compared, identifying best and promising practices is not possible” (p. 6). A systematic review of this literature allowed for the collection of materials from a variety of sources across the country and creating a manageable starting point for future dialogue.

ESTABLISHING CONSISTENCY OF FINDINGS

The advantage of looking at both qualitative and quantitative elements within a series of studies about a particular topic is that the individual studies often take place in different contexts and, in the case of qualitative research specifically are intricately tied to the researchers experience and particular worldview. A single study can rarely provide a definitive answer to a research question focused within the social sciences (Cooper, 1984; McGaw, 1997). The educational research environment is often difficult to control (even if

you wanted to) and human behaviour complex to explain (Wolf, 1986). In addition, economic constraints particularly in fields like education may restrict the scope and scale of any single study (Suri, 2000). Synthesizing primary studies can improve our understanding of not just what occurs but how and why.

INTEGRATING CRITICAL INFORMATION FOR DECISION-MAKING

In the field of education the potential of systematic reviews to provide policy makers, teachers, parents, and students with critical information from which to design instruction, structure school environments, develop teacher professional activities, etc. cannot be underestimated. Dropout prevention programs have reported mixed results in terms of their success at reducing dropout rates among learners, by integrating and summarizing what the available and relevant literature reports could help decision-makers make more informed decisions about how and when to allocate resources, how to modify or implement programs and also draw attention to specific implementation considerations, particularly when programs are targeting a specific sub-group of learners.

COMPARING THE INFLUENCE OF RESEARCH DESIGN

When a policy maker, teacher or parent is faced with making a decision, can we be confident that that research design provided the most comprehensive review of the phenomenon in study? Knowing that any single study is almost certain to have used a single research design; can we make decisions and draw conclusions about a study based on a research design that may include specific constraints, biases or limitations? Even a well-done study may not foster great confidence. This is where systematic reviews can help. It allows us to compare the research designs that lead to specific findings.

ENCOURAGING THEORY BUILDING

A final purpose for conducting a systematic review of this literature relates to its potential contribution to theory development (Estabrooks, Field & Morse, 1994). In effect, meta-level analysis allows us to explore the validity of theoretical relationships between variables within the instructional context. While not every synthesis is able to provide definitive theoretical evidence, at the very least they can provoke theoretical discussion and help direct future research (Cooper & Hedges, 1994, p. 21).

Limitations and Challenges in Conducting Systematic Reviews

Systematic reviews are not without their own methodological limitations and challenges. Cooper and Hedges (1994) would refer to these kinds of concerns as threats to the validity of the synthesis. Validity however is not a term used widely in qualitative literature. The term “trustworthiness” has gained increasing currency as a useful way of thinking about the confirmability, credibility and transferability of qualitative findings (Bair, 1999; Lincoln & Guba, 1985). In attempting to incorporate different kinds of primary data in this proposed study, it is important to consider constructs that relate both to validity and to trustworthiness as they relate to systematic reviews. The full extent of these potential limitations and issues are addressed in detail in places like *The Handbook of Research Synthesis* (1994) and the EPPI-Centre systematic review protocols. Many of the methodological concerns associated with qualitative synthesis are concerns we have about any type of secondary analysis, quantitative or qualitative (Suri, 2000; McGaw, 1997) and thus the following sections highlight key methodological considerations for this study.

THE QUALITY AND INTERPRETATIONS OF PRIMARY STUDIES

The quality of a systematic review is completely dependent on the quality of the primary studies selected for inclusion (Bair, 1999). Limited space for reporting results is always a concern with meta-level research because of its potential for restricting the publication of information that is critical for assessing the quality of primary studies.

When researchers are forced to condense or leave out portions of the analysis for the sake of space the question of how research bias will affect synthesis is a serious concern (Cooper, 1984). A further concern about primary studies has to do with what Mullen and Rosenthal call the “file drawer” problem (1985). They estimated that only 5 percent of studies with significant results are published and 95 percent of studies have no significant results and are thus relegated to life in a filing cabinet (Bair, 1999). This issue is further compounded by growing concern over the reliance on relatively underdeveloped electronic databases as the primary source for locating studies (Oakley, 2003). Oakley suggests that many databases “lack comprehensive keywording, any thesaurus of standardized search terms or sophisticated search strategies, and each requires individualized approaches” (p. 27). Finally the quality of the primary studies themselves, their research questions, research designs and ability to report results that are clearly rooted in the data are beyond what the meta-researcher can control for. Cooper (1984) asserted that “the unreliability in data introduced by the incomplete reporting of primary researchers” (p.77) is a critical question of credibility and trustworthiness. It is critical then that through the synthesis of primary studies, the meta-researcher identifies those concerns and makes credible decisions about whether the study should be included at all.

LACK OF COMMON LANGUAGE AND REPORTING STYLES

The debate over whether one should include or exclude studies that used a different methodology will likely never be decided. Researchers will continue to have to address and describe how the phenomenon and the findings of a systematic review that excludes alternative research designs (e.g., only examines randomized control trials) could have been influenced by research design. Cooper (1984) contended that when we code study methodologies we should "exhaust as many design moderators as possible" (p. 77).

Apart from that, there are deeper issues related to synthesizing different researchers' interpretations, linguistics styles, and writing conventions. Light and Pillemer, (1984) claimed that by virtue of their emphasis on idiographic knowledge and the complexities and contradictions inherent in lived experience, qualitative studies may resist our efforts at 'summing up'. Sandelowski et al. (1997) highlighted this concern saying "it seems both epistemologically and ethically inappropriate to attempt to summarize findings from one or more qualitative studies about human experiences" (p. 366). Each primary study was written, interpreted and conceptualized by an individual with a particular philosophical, theoretical perspective. Can we translate the metaphors, interpretations of one researcher into another? In summarizing do we ultimately lose the energy, distinctiveness and vividness of the human experiences represented in the primary studies?

Sandelowski et al. (1997) made another interesting point in saying that systematic reviews involving qualitative studies require the researcher to engage in the interpretation of culturally diverse texts. This view eschews notions of knowledge building from a collection of evidence and is distinctly post-positivist and radical constructivism in its form. Accordingly, a collection of studies cannot be synthesized because the

generalizability of findings, across people, settings, and variables is an unattainable and undesirable goal. Other researchers, policy-makers and practitioners reject this perceived “radical” view with a notion of science and the firm belief that truth may be collectively known.

Despite these concerns there is a growing effort to construct new frameworks for including different kinds of evidence in systematic reviews (e.g., Oliver, Harden, Rees, Shepherd, Brunton, Garcia & Oakley, 2005). As systematic reviews are becoming increasingly used as tools to support policy making, reviewers are looking at ways of combining evidence from qualitative and quantitative studies in order to capture the full complexity of program interventions, its impact and its transferability to other contexts. Where quantitative studies help us answer questions of whether an intervention works, qualitative methods “provide insights into why and how complex initiative work” (Boaz, Ashby & Young, 2002, p. 10). Researchers then will likely have to find ways of analyzing the studies for features like rhetorical devices, plot lines, semantic and idiomatic translation (Noblit & Hare 1988; Sandelowski et al., 1997, p. 370).

Beyond just what the studies reveal the reviewer must also consider how the researcher reported it. Given the wide variety of writing styles and presentations styles, particularly for disseminating qualitative research, any attempt to blur those lines may be seen as an attempt to undermine the philosophical, social, political and ethical commitments of the primary authors. These issues must be thoughtfully considered and evaluated and care must be taken to maintain the integrity of individual studies.

ACHIEVING VALID CODING OF PRIMARY STUDIES

Achieving reliable coding of study features is also a concern with meta-analysis and in fact many forms of integrative reviews (Jackson, 1980). When dealing with a small number of studies, it is more likely that a single investigator will conduct the coding. The nature of many qualitative studies in particular is thick descriptive and interpretive writing and if the demands of coding such thick writing are spread over a number of days and weeks it may “raise serious threats to coding stability” (Jackson, 1980; p. 454). Furthermore as the number of studies in a systematic review increases we may want to involve other coders to help with the volume of studies. Inter-coder agreement becomes a priority. More than the logistics of coding, researchers must also be concerned about the actual codes. Cooper (1984) argued that one of the major threats to validity (and also trustworthiness) is the “unreliability in coding the research results” (p. 77). Unreliability in the coding process adds random variation to the observations, increasing estimates of standard error and potentially affecting interpretations of correlations among study features. When the meta-researcher can retrieve information directly presented by the primary researcher this is not as much of a concern. When the meta-researcher makes inferences about the research quality or the presence or absence of particular concerns over credibility, however; the validity of the synthesis may be questionable (Cooper, 1984). The phenomenon under investigation may involve very different constructs, terminology and language. The researcher will have to develop a coding scheme that accounts for these subtle and not so subtle differences.

POLITICAL CHALLENGES

Like many forms of educational research that bears some component of evaluation, the use of systematic reviews like the one described in this proposal is taking place within a disputed context. Oakley (2003) asserts that:

“there is a sizeable voice in the educational research world about.... limited questions about ‘what works’, outdated notions about the role of ‘procedural objectivity’ and a disregard for the tenants of postmodernism which question the validity of any a priori framework of enquiry” (p. 26).

In fact some do not support the current use of research synthesis arguing that in the majority of cases where it is being used it “muddies the waters, disregards the problems and leads to meaningless conclusions that are likely to hamper proper scientific research” (Eysenck, 1984, p.58). Educational issues like dropout prevention that are “hot topics” within the media and national educational discourse further compound the political complexities of conducting systematic reviews. The researcher conducting systematic reviews must not only consider the methodological and practical implications of their study designs but the political contexts in which they seek to inform, direct and influence educational policy and practice

Locating Available Program Documentation

Searching for studies in a systematic review is a critical as it circumscribes the population of studies that will be examined. It is important to develop the most appropriate search in order to accurately represent the field of study and locate as many relevant studies as possible. According to Rothstein, Sutton & Borenstien (2005) “... if the sample of studies retrieved for a review is biased than the validity of the results... no matter how systematic and thorough in other respects, is suspect” (p. 2). A recent study by McLeod and Weisz (2004) compared effect sizes of dissertations and published

studies of child and adolescent psychotherapy in order to investigate publication bias. The findings of that study demonstrated that dissertations reported effect sizes with less than half the magnitude of those reported in the meta-analysis of the published literature. A recent review for the Campbell Collaboration (i.e., Nye, Turner & Schwartz, 2005) found similar results. With that in mind the process for retrieving relevant studies was conducted with attention to minimizing bias and to consider studies beyond those in the published literature.

A comprehensive mapping of potential search terms and channels for locating relevant studies was conducted and revealed five key areas for study identification. Identification of studies for inclusion was conducted by applying selection criteria adapted from previous and related reviews as well as guidelines presented by recognized research centres.

IDENTIFYING RELEVANT SEARCH TERMINOLOGY

Drawing on existing procedures for synthesizing both qualitative and quantitative studies (e.g., Hossler & Scalese-Love, 1989) the initial stage of the project focused on the identification of both qualitative and quantitative studies through a comprehensive search of publicly available literature between 1990 and 2006. The idiosyncratic use of language and method in the field of dropout prevention necessitates using literature retrieval procedures that are sensitive to cultural differences and linguistics subtleties (Thorne, 1994). Pettigrew & Roberts (2006) maintain that all search strategies are basically similar in that they make use of key outcome and intervention synonyms –all combined with the AND, OR or NOT connectors. Thus once all possible dropout, prevention, outcome, setting, and age terms were identified using database thesauruses and key conceptual

resources they were combined in various ways with each other to extract as many relevant studies as possible from the literature. Search terms were used differently for some sources but the main terms used would include “dropout prevention programs”; “dropout intervention” “student retention programs” and “improved graduation rates”: To further ensure that relevant studies are not excluded a series of related French terms were also be used including for example “l’abandon scolaire; “ décrocheurs” and “ l’intervention différentielle”.

1. *Dropout terms included:* dropouts, early-school leavers, non-graduates, pushouts, disaffiliates, educational mortalities, stopouts, at-risk, potential dropouts, out of school youth (Weber, 1987)
2. *Intervention terms included:* programs, prevention, interventions, initiatives, services, early-intervention, re-entry, work transition, educational programs, support services, flexible hours, individual instruction, dropout rehabilitation, high-school equivalency programs
3. *Outcome terms included:* retention rates, graduation rates, academic achievement, education attainment level, diploma, school graduation, school expulsion, dropout rates, school retention, truancy, persistence, student attrition, GED, outcomes of education, treatment outcomes, outcomes of treatment, school to work, transition, school -to -work transition, school transition
4. *Setting terms included:* schools, care facility, accelerated programs, alternative education, non-traditional education, alternative programs, alternative schools, correctional institutions, high schools, middle schools, secondary education, elementary schools, pre-school, early childhood, junior high schools, technical school, vocational school, vocational education, vocational high school, storefront schools, outreach schools

DEVELOPING SEARCH STRATEGIES

Due to the challenges posed by attempting to search comprehensively, persons interested in conducting systematic reviews are encouraged to seek the assistance of those with expert skills in information retrieval strategies (e.g., White, 1994). With this in mind, Anne Wade (Centre for the Study of Learning and Performance) was consulted for advice regarding the planned sources of studies as well as for advice regarding the development of advanced search strategies. While primary research involves a relatively well-framed procedure for collecting data, research synthesists must consider multiple channels for accessing literature in order to reduce publication bias and include as many relevant studies as possible. To this end, this project used five main channels for identifying relevant literature.

1. **Studies addressing dropout prevention programs in databases and select journals.** While there is no hard and fast rule about the number of electronic database searches required to produce a high-quality systematic review a few recommendations have been suggested by meta-analysts. Petticrew & Roberts (2006) recommend a minimum of two databases, although they note that this will vary considerably with the review topic. Lipsey & Wilson (2001) suggest a general “multiple databases” approach. Hopewell, Clark and Mallett (2005) advocate searching multiple types of database in order to tap into the unpublished literature. Petticrew and Roberts (2005) prefer a sensible approach, arguing that time and available resources ultimately dictate the number of sources searched. In an attempt to address these challenges of locating published material, several computerized searches of databases were conducted including:

Educational Resources Information Center (ERIC), Proquest Digital Dissertation, PsychInfo, EBSCO Academic Search Primer, ABI Inform, CBCA, Educational Technology Abstracts and the Canada Research Index. The Library and Archives Canada database was also searched. The abstracts of journals including the Canadian Journal of Program Evaluation Journal of Education for Students Placed at Risk were hand searched for potential articles (1990-2006) and they were filtered to identify those reporting on Canadian programs. The bibliographies of each article in the special issue of Exceptionality Education Canada (vol. 13) on Students At-risk in Canadian Schools and Communities were also reviewed for potential documents (Examples of specific search strategies can be found in Appendix A).

2. **Direct contact with governmental departments, community organizations and research agencies.** Informal channels of communication which do *not* require explicit rules on the part of a primary researcher for gaining access to that channel (e.g., published articles in a journal might require peer-review) play an important role in accessing the grey literature. It was expected that the majority of reports would come directly from governmental archives, local school boards and community organizations and over 200 contacts were made via email and telephone to identify potentially relevant documents for this review. Requests for information related to ongoing or completed (and published or unpublished) projects. However; few agencies, program administrators, school boards etc. contacted either had available reports or would make those reports publicly available. Each provincial department of education was contacted for, annual

reports, task force reports, policy documents, program evaluations, research documents, etc. that report directly on dropout prevention programs in their region. The search also involved contacting individual program administrators within the community and school boards who were identified through branching (e.g., The Niagara Peninsula Industry Education Council). Private companies which publicize their private funding for dropout prevention programs (e.g., Royal Bank of Canada “Stay in School” project and Aboriginal Stay in School program) were also contacted. In addition national and regional research centers and organizations (e.g., Centre for Research on Youth at Risk, Canadian Education Association, Society For the Advancement of Excellence in Education) were also contacted and/or websites reviewed to further identify and collect potential documents for review).

3. **Web-based search for documents on dropout prevention programs in Canada.** A Web-based search for dropout intervention and prevention programs was conducted. The search terms described above were employed to search government websites, research databases (e.g., Canadian Research Index) as well as free text searches using the widely available Google search engine. Conference proceedings (e.g., The Canadian Evaluation Society, Canadian Education Association) were also searched online. Targeted web-based searches were conducted, for example: (a) names of researchers who had published a potential study for inclusion were searched to see if they had published additional research (e.g. Volpe); (b) names of programs which were reported in newspaper articles or trade journals which themselves did not meet inclusion criteria were searched via

the web to locate potential reports (e.g., Pathways in Toronto); and (c) documents identified through branching which did not indicate where the report was located or available were also searched for using targeted web-based searches

4. **Contacting researchers in Canada working in the field of dropout prevention and youth at-risk.** Canadian-based researchers working in the field of youth at-risk and dropout prevention programs were identified through Canadian researcher networks (e.g., The Learning Partnership (TLP), Centre For Research On Youth At Risk- New Brunswick, Atlantic Alliance for Youth, National Youth in Care Network, La Fédération de la jeunesse canadienne-française (FJCF), The Canadian Adolescents At Risk Research Network (CAARRN), etc.) as well as through the Government of Canada supported Centre of Excellence for Youth Engagement. In addition, authors who are identified through retrieved documents were also be contacted if there was reason to believe they had published relevant documents. Sample letters (English and French) can be viewed in Appendix D.
5. **Branching from other sources.** A review of references cited in other articles, documents or reports that did not turn up through other search strategies were conducted. In addition the dropout bibliography put out by the Ontario Secondary Schools Teacher Federation (2004) was reviewed.

Inclusion /Exclusion Criteria

There are many considerations that relate to setting criteria for inclusion and literature reviews on the topic of dropout prevention programs (e.g. Prevatt, 2003) have consistently noted several key limitations of working with this literature; (a) the lack of common accepted definition of “dropout”; (b) the concentration on predictor variables

(e.g., peer relationships, grade retention, stress and coping, persistence, socio-economic status); (c) the inability of the current literature to capture the complexity and magnitude of the issue because of the high number of dropout related correlates (i.e., Rumberger 1998, identified 100 non-school variables which correlate to dropping out); (d) the difficulty of using random assignment research design when evaluating school based programming; and (e) definitions of school success broadening to include not only graduation statistics but social and behavioural standards (Christenson, Sinclair, Lehr and Hurley, 2000).

Given these challenges it was conceivable that this project could have ended up reviewing materials in a vast number of related but distinct literatures including, educational programming for literacy, educational programming for students with special needs, generalized curricular reform, school climate, counselling services for families and parents among others. However, as noted by Lehr (2003); “the extent to which these interventions are systematically targeted for disengaged learners is unclear” (p. 343) and for us to draw conclusions about which of those programs could be conceived of as ‘dropout prevention or intervention’ would have been irresponsible. I support the actions taken by local (e.g., Partnership Table for School Retention in Montreal) and national (e.g., Engagement and Dropping out a Life Course Perspective –HRDC) organizations who argue for a better understanding of the complex interplay between individual, family, peers, school, and community factors relation to dropping out. However, the scope of this present review was not to examine how educational programming in general supports school completion but to explore how programs in Canada that self-identify as being a dropout prevention program have designed and implemented over the past two

decades and the kinds of success and challenges they have observed. The broad and inclusive search terms and searches strategies developed were designed to that end.

With those issues in mind the inclusion criteria were operationalized to provide a context for the question and framing how studies were examined. As an initial step in determining documents that might be included in the review a preliminary set of inclusion criteria was used, studies identified as potential includes would:

- Involve the influence of an educational or community organization in the planning and preparation of the program (this distinguishes it from personal enrichment).
- Explicitly identified the program as focusing on dropout prevention or intervention
- Provided either impact data (qualitative or quantitative) on program effectiveness
- Be publically available or archived (or will be archived)
- Have been published between 1990 and 2006
- Be written in either English or French.

After a list of potential references was generated from the search strategies, the titles and abstracts were scanned to determine if studies appeared to meet the initial criteria. When the title and abstract were not comprehensive enough to make a decision the document was obtained to identify pertinent information for making inclusion decisions. Following procedures used by Lehr (2003) in a related study, any studies identified as “questionable” during the initial scanning of abstracts process was by included in the next stage of selection procedures and evaluated using more detailed inclusion/exclusion criteria. Documents rejected for inclusion during the initial stage were first scanned to determine if any additional references could be located using branching.

The second stage of selection procedures involved developing and refining a more specific inclusion/exclusion criteria using similar techniques as previous reviews (e.g.,

Bernard et al., 2004) as methodological questions emerged after examining the initial pool of documents. Comparisons with inclusion criteria from other related systematic reviews (e.g., Prevatt & Kelly, 2003; Lehr, Hansen, Sinclair & Christenson, 2003) was also conducted to explore additional issues for consideration. The following set of exclusion criteria were used:

Excluded from the review are:

- Studies describing a work/school transition program that is targeted at students who have already completed high school.
- Single participant studies (e.g. excluded study: Efron, 1990)
- Studies which describe a proposed program or model intervention that has not been implemented (e.g. excluded study: Milak, 2001; Phaneuf, 1993)
- Studies which report on a program that targeted individuals who had dropped out of school but was not aimed at them completing high school (e.g., Currie, 2001)
- Studies and reports which describe several unique and varied programs with no way of separating individual program outcomes from included initiatives (e.g., Gordon, 2003)
- Studies which describe a program but do not provide impact data, either qualitative or quantitative related to desired outcomes (e.g., Ford, 1999)
- Studies which describe a program for at-risk learners which are not explicitly described as a dropout prevention program (e.g., Barron, 1995)
- Studies which use data provided in another study/report accepted for inclusion (e.g., Coutts et. al., 1995)

Developing Coding Scheme

Given that the majority of primary studies on dropout prevention programs traditionally use non-experimental case study designs and come from a variety of sources (e.g., peer-reviewed journals, technical reports, governmental briefs, etc.) it was critical to develop a coding scheme which allows for extracting information on a variety of

methodological and conceptual features of the studies under review. The project drew on the guidelines established by national organizations for systematic reviews (e.g., EPPI, International Campbell Collaboration and What Works Clearinghouse) as well as the coding schemes used in related reviews (e.g., Lehr et al., 2004; Aron and Zweig, 2003) and published methodological approaches for incorporating qualitative data in systematic reviews.

Table 4. Guidelines for Systematic Reviews

Coding Manual/Review Guidelines	Source
The Procedural and Coding Manual for Identification of Evidence-based Interventions	Kratochwill et al. (2002)
Evidence for Policy and Practice Information and Co-ordinating Centre (EPPI-Centre)	http://eppi.ioe.ac.uk/EPPIWeb/home.aspx
What Works Clearinghouse (WWC)	http://www.whatworks.ed.gov/
Using Qualitative Studies in Systematic Reviews	Dixon-Woods et al. (2006)
Conducting Meta-study of Qualitative Research	Paterson et al. (2001)

The coding scheme is designed to provide a summary of (a) the range of programs and approaches reflected in the literature; (b) examine the extent to which studies reflect contemporary constructs related to dropout/school retention (e.g., intervention focus, outcome measures); and (c) critically investigate the degree to which studies report on threats to validity (e.g., sample size, statistical significance, use of outside evaluator). Stock (1994) cautioned that creating extremely extensive coding schemes could create problems in that (a) as the number of study characteristics to be coded increase so too does the length of time required to code each study, this in turn can impact on the

researchers ability to analyze the results in a timely manner and (b) adding items increases the probability of reporting at least one chance relation as significant. To address this issue Stock (1994) proposes that the researcher formulate a coding scheme which is based on reasonable conjectures about the relationship between effect sizes and study characteristics. Conjecture here being defined as both first hand knowledge of the research domain and theoretical evidence (p. 126).

An initial list of codes for extracting data from included studies can be seen in Section 9 of the preliminary codebook (Appendix B). Specifically the codes for this project are divided into the following categories:

- study identification (e.g., author, year, source of publication)
- methodological features (e.g., research design, effect sizes, durability of treatment, implementation fidelity, sample size, etc.);
- objectives (e.g., program objectives, student population targeted elaborate?, etc.)
- environmental features (e.g., urban/rural, age group, linguistic or racial composition of students, etc.)
- resources (e.g., program funding source, support personnel, costs benefit, training, etc.);
- components (e.g., use of technology, degree of self-paced instruction, student/teacher ratio, types of assessment, degree of mentoring or peer support, etc.)
- management (e.g., use of external evaluator, type of program, duration of intervention, etc.)

Since the focus of this review was on synthesizing existing evidence on the effectiveness of educational interventions and not on documenting the complex process of early school leaving as a social, the coding scheme was designed specifically to capture as much information as possible about why programs were designed and implemented the way they were (i.e., was there a particular predictive factor related to dropout the program was trying to address?; was there a particular instructional philosophy being applied?; were

there distinguishing demographic features about the students participating in the intervention?) Thus, while the review did not intend to explore *why* students drop out generally, it did attempt to capture how individual interventions defined and expressed their role in addressing various risk factors relating to dropout and how that may have related to measures and outcomes of effectiveness.

Quantitative Methods for Extracting Data

Where applicable, effect size calculations for studies in this review were computed following procedures outlined by Glass et al. (1981) with corrections for sample size (Hedges & Olkin, 1985) being followed. The effect size is a statistical representation of the magnitude of the relationship between two variables. Statistical procedures standardize the data from each individual study. The standardized data are reported as an effect size. The Effect size (ES) is the standardized mean difference between the intervention and control group, that is $ES = (X_e - X_c) / SD_p$ where X_e and X_c are the intervention and control group's means, respectively, and SD_p is the pooled within-groups standard deviation. Comparison of separate study effects is made possible through the use of effect sizes. Differences in sample size were encountered both within and across multiple studies. As effect size estimates for larger sample sizes are more exact than those based on smaller sample sizes, weighting of effect size estimates based on the sample size is required. The effect size is multiplied by the inverse of its variance (proportional to sample size) to obtain the weighted effect size (Hedges & Olkin, 1985). Comprehensive Meta-analysis (CMA) version 2.2.021 (Biostat) was used to calculate the

weighted effect sizes. This software has several notable features, most useful for this project is the ability to accept data in different formats and transform it to a common effect size and variance. Because results have been transformed to a common metric, the magnitude from different studies can be compared.

In instances where data were provided in the form of t tests, F-tests, p levels and frequencies, effect sizes were computed following the appropriate conversion formula detailed by Glass et al. (1981) or Hedges Shymansky and Woodworth (1989). Formulas can be found in Appendix C. During coding procedures these findings were coded as “estimated effect sizes”. The unit of analysis was the independent study finding; thus multiple outcomes could conceivably have been extracted from the same study. (e.g., the study reports changes in dropout rates as well as student attitudes towards school). Rules governing the calculation of effect sizes emerged as the data were retrieved and coded. Examples of initial rules were: (a) impact or dropout rate when multiple achievement data are reported (e.g., initial dropout rate, percentage of learners going on to college, percentage who eventually graduate from high school) dropout scores were used in calculating effect size; (b) effectiveness or satisfaction rate for attitude inventories, the average of all items failing under one type of outcome (e.g., attitude toward school) were used so that only one effect size is generated from each study for each outcome. This rule only came into effect for one study which reported multiple academic outcomes (i.e. reading, writing and mathematics scores) (Saint Laurent, 1996), and c) in cases where a single comparison group was compared to multiple treatment groups (e.g., Mutadi, 1990) we calculated individual ES for each treatment but used N/3 for the control group sample size.

Qualitative Methods for Extracting Data

To cull additional qualitative data from the studies being reviewed in this proposed project, a process outlined by Fitzgerald (1995) was used as a guide for creating almost a second layer of coding. Fitzgerald proposes a systematic interpretive procedure that closely resembles a constant comparative method often found in primary studies and many of these steps are simply a complementary and more detailed level of coding the study features described above: For example Fitzgerald suggested that:

1. In addition to the quantitative data elements collected in a typical meta-analysis, a combination of detailed descriptive notes regarding features of the study (i.e., participants, treatment, instruments, etc.) and categorizing of information into detailed codes be conducted
2. These detailed notes are reviewed to identify themes, contradictions, patterns or other issues needing to be addressed by the researcher. A series of topic “clusters” are given tentative labels. As an example, while “parental involvement” was coded for in the standard study features coding, detailed qualitative notes which described unique or particular details about the kinds of parental involvement being described by the author or particular implementation issues or challenges brought forward in the study so one could potentially go back and explore similarities and difference within that subset of studies.
3. Study features and corresponding notes are reviewed and individual studies are sorted into tentative clusters (it is possible for a study to fall into multiple clusters). Clusters of studies in SPSS were examined and re-examine of notes on studies within those clusters was conducted to provide additional context to the numeric or categorical data.

4. The similarities and differences among studies within each cluster would then be examined. At this stage, Fitzgerald proposes generating initial hypothesis from summarizing and sorting meaningful results and then re-examining studies to review discrepancies and resolve potential contradictions.

Assessing Research Quality

Given that this proposed project involves collecting data from research studies employing multiple forms of research designs it is important to develop means for assessing the quality of the primary studies in order to frame the review findings. In general research quality refers to the “accuracy” of the research findings (Lecompte & Goetz, 1982, p. 32). As with the integration and analysis of the research findings the criteria to assess research quality was dependent on the types of studies retrieved and the available data extracted.

The standard criteria for assessing the validity/trustworthiness of studies was derived from the LeCompte and Goetz (1982) text which translated Cook and Campbell’s (1979) work on threats to internal and external validity that exist in research studies employing qualitative research designs. According to Lecompte and Goetz “establishing validity requires determining the extent to which conclusions effectively represent empirical reality and assessing whether constructs devised by the researchers represent or measure the categories of human experience that occur: (p. 32). The criteria for establishing validity outlined by Lecompte and Goetz (1982) were utilized to assess the research quality of the studies selected for review.

Based on these criteria a four point scale was developed to assess the validity of each outcome. Table 6 outlines the criteria. These scores were assigned during the data extraction process and were used during analysis to explore potential patterns or inconsistencies in findings based on level of research quality. The scores may also be used to highlight current trends in research design in the field and provide direction for future research. For quantitative studies a validity rating of *high* was assigned when a true experimental design with random assignment was utilized. A rating of *moderate to high* was assigned to non-equivalent pretest/posttest designs with appropriate statistical procedures and evidence of controlling for possible threats. A non-equivalent pre-test/post-test design with few controls was rated as *moderate to low*. One-group pre-test post-test, non equivalent groups pretest only, and one-group post-test only designs were rated as having *low* validity.

For qualitative studies a rating of high validity/trustworthiness was given when studies corroborate evidence through triangulated methods of data collected, research design and data collection are clearly described and where personal and intellectual biases are explicitly stated. A rating of moderate to high trustworthiness/validity was given to studies which provide detailed descriptions of the research process, present consistency between theoretical orientation and types of data collection and provide some evidence to corroborate findings. A rating of moderate to low was given to studies which provide limited corroborating evidence and description of research process and theoretical orientation. Studies which do not triangulate data, present no specific theoretical or methodological orientation and which present a very limited description of research procedures was rated as low trustworthiness/validity.

Table 5. Criteria for Assessing Research Quality

Validity rating	Brief Description
<i>Quantitative</i>	
High	True experimental random assignment
Moderate to high	Non-equivalent pretest/posttest designs with appropriate statistical procedures and evidence of controlling for possible threats.
Moderate to low	Non-equivalent pre-test/post-test design with few controls
Low	One-group pre-test/post-test, non equivalent groups pretest only, and one-group post-test only
<i>Qualitative</i>	
High	Data triangulation (member checks if appropriate); clearly described research procedures; personal and intellectual biases explicitly stated, discussion related to negative cases or contradictory findings.
Moderate to high	Detailed descriptions of the research process; consistency between theoretical orientation and types of data collection and presentation of corroborating evidence.
Moderate to low	Limited presentation/discussion of corroborating evidence; vague descriptions of research process and theoretical orientation
Low	No identifiable theoretical or methodological orientation used very limited description of research procedures, no attempts to triangulate data.

Results

Outcome of Searches

For the Canadian studies, approximately 987 database abstracts and executive summaries concerning dropout prevention programs were searched and 134 full text items to be retrieved were identified. Through Web-based searching an additional 61 potential includes were identified. Thirty-seven reports and documents were forwarded from government and research contacts and an additional 8 potential includes were identified through branching. In addition to these studies, there were a significant number of program reports/evaluations initially identified which could not be located and/or retrieved. These non-retrievable documents were normally: a) school board documents (e.g., *Rampal Turner, C. (1994). Programs and activities at the Scarborough Board that*

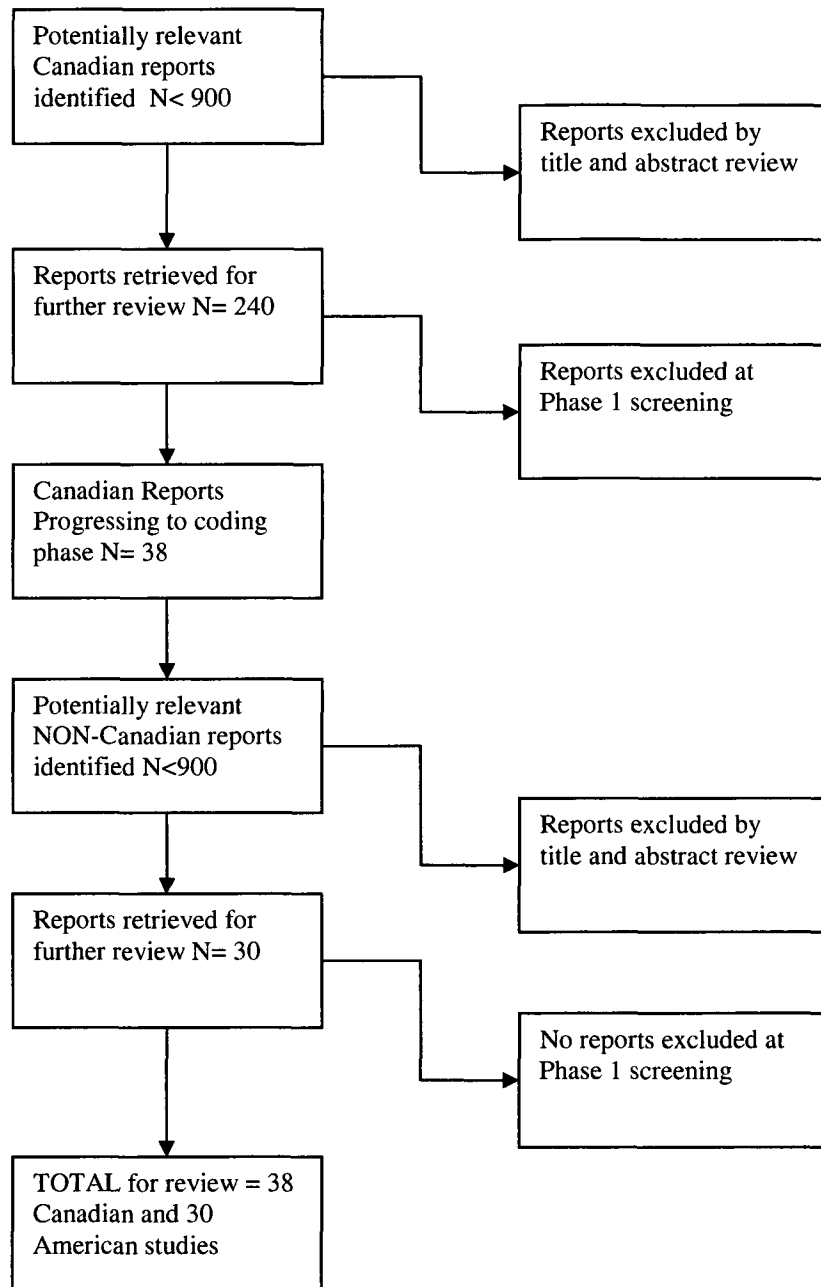
are aimed at reducing the dropout rate (#93/94-27). Scarborough, ON: Scarborough Board of Education. *Evaluation of the Outreach Program Overview Report, 1990*); b) documents held by organizations that do make the information publically available (e.g., *Parsons, T. E. (1993). Beyond School: A Stay-in-school Program for Grade 9 Students at Risk of Dropping Out, 1993 (371.2913 P271) Alberta Teachers Association.*); or c) were documents where not enough bibliographic information was provided to locate relevant documents. (e.g., *Kingston: The Special Delivery Club*).

In total 240 documents were reviewed for inclusion and from them 38 met all of the inclusion criteria. Figure 1 summarizes the formal search and retrieval process. A sample of 10 studies was reviewed by two independent raters for possible inclusion and to test for reliability of the coding methodology. Ideally inter-rater agreement would have been calculated on the complete set of studies at all steps in the coding process however, given budgetary and time concerns that was not possible and being able to assess a degree of reliability at different stages of the coding was deemed more prudent than exhausting those resources for only one dimension. Several steps, as suggested by Matt & Cook (1994), were taken, however, to ensure that the rating comparisons which were undertaken were rigorous. A detailed codebook was used by both coders that included the purpose of the research, the research question and instructions on the inclusion/exclusion criteria for evaluating potential studies along with detailed descriptions of study feature codes, including examples of how those features might be “evidenced” in the literature (Appendix B). In addition, both coders had experience in applying meta-analytic procedures and in using protocols and criteria to inform inclusion decisions, study feature coding and effect size extraction in this type of review. Both coders were familiar with

procedures on using the particular set of excel spreadsheets for recording coding decisions.

A percentage agreement method (Morgan, Gliner & Harmon, 2006) was used to evaluate the level of agreement between raters. The initial inter-rater agreement as to inclusion was 83.3%. Coders met to come to a consensus decision on studies for which disagreements had occurred. After this meeting, there was agreement among coders.

Figure 1. Overview of Selection Process For Including Reports



There were two main reasons studies were excluded: (a) studies were conceptual articles describing issues related to dropout prevention; and (b) studies violated exclusion criteria related to definition of dropout prevention program. Table 6 provides an overview of exclusion decisions.

Table 6. Categories, Numbers and Percentages of Excluded Canadian Studies

Category	<u>Excluded Studies</u>	
	Number	%
Conceptual paper or report	104	52
Violated “dropout program” definition	79	39
Did not report on program outcomes	13	6
Duplicates	6	3

In addition, a comparison of group studies conducted outside of Canada were also retrieved. After the Canadian studies were identified for inclusion a review of how those studies had been located was undertaken. Sixty percent of the Canadian includes were identified via free text Web-based searches or through branching and the remainder through traditional database searches. In order to match the sampling strategy for the 30 non-Canadian studies, a random sample of equal proportion of documents via the Web using the table of random numbers from Eckhardt & Ermann, (1977) and an equal proportion of studies randomly from ERIC. Although the sampling strategy sought only to identify studies conducted outside Canada, no studies outside of North America were identified using the sampling strategy, thus in the end a group of 38 Canadian studies and a group of 30 American studies were included for review.

Coding of Studies

The literature retrieval and coding process yielded 151 individual interventions that were described in 68 written reports. Some reports (e.g. Stay in You Win, Hayward) reported on a large number of independently run programs in the same report. In addition, many reports provided data on multiple student groups over the course of a few years (i.e., independent cohorts). As outlined in Table 7, only nine of the 68 reports were published after 2002 and most were published as journal or trade journal articles. Each program was coded using the codebook (see Appendix B) and codes were entered into an

Excel Spreadsheet, and later, SPSS, for analysis. Inter-rater agreement was calculated for a sample of 10 studies (including both French and English reports). Initial inter-rater agreement rates for study feature coding was 84% with complete agreement after coders met to discuss discrepancies. Disagreements may have been explained in part by multiple category response items which did not use “YES”, “NO” categories but had a variety of potential responses. This increased the potential for divergent evaluations within individual items.

Table 7. Publication Information for Sixty-Eight Written Reports

Publication Date	Canadian		Outside of Canada	
	Frequency	%	Frequency	%
1990-1995	15	40%	13	43%
1996-2001	19	50%	12	40%
2002-2006	4	10%	5	17%
Publication Category				
Journal Article	19	50%	7	23%
Dissertation/Thesis	5	13%	2	7%
Book Chapter	1	3%	0	0%
Technical reports	13	34%	21	70%
Validity				
High	0	0%	5	17%
Moderate to high	5	13%	5	17%
Moderate to low	11	29%	12	40%
Low	22	58%	8	26%

Each report was also read for familiarization with any additional qualitative data reported. Identification of major themes emerging from each report was conducted using an open-ended coding scheme and standardized reporting format. Coding was dynamic and recursive rather than sequential, meaning that each study was reviewed multiple

times so that themes and premises could be understood in relation to other studies. Clusters of studies were then juxtaposed, cross-compared and integrated by asking questions like Were there similarities, contradictions or commonalities across studies relating to this theme/position/argument?; and, If you only took the interpretation of one study how would the “picture” of the theme be incomplete?; to develop a refined and comprehensive description of the key qualitative data relating to dropout prevention initiatives.

Analysis of Effect Sizes

The intention of the project was to calculate effect sizes in order to provide additional analysis of intervention impacts, however 47 out of 68 reports reviewed in this project did not report effect sizes or power analysis, or provide sufficient numerical data on program impacts to extract that information. Of the thirty-eight Canadian reports and studies reviewed only 8 (21%) of them provided the necessary information for effect size calculations. Of the 30 American reports reviewed 13 (43%) provided that information.

If impact variables were not measured (e.g., dropout rates, absenteeism, graduation rates) effect sizes were calculated for other dependent variables to explore more broadly the program impacts being evaluated (e.g., attitude toward school, behaviour rating). In cases where multiple outcomes were reported ES, multiple ES were then extracted. Inter-rater agreement was calculated for effect size computations using a sub-set of 12 reports using the same calculation method as with exclusions/exclusion agreement. Agreement rates were 99% on the number of effect sizes that could be extracted and 100% for effect size calculations themselves.

In total, 66 effect sizes were extracted (See Appendix E) from 21 different reports. Interpretation of the effect sizes reported for each study applied guidelines presented in Lipsey and Wilson (2001). Additional technical resources were used to determine the best approach for defining, validating and summarizing these reports (Hedges, 1981; Hunter & Schmidt, 2004).

Differences in sample size were encountered both within the same study and across multiple studies. As effect sizes estimates for larger sample sizes are more exact than those based on smaller sample sizes, weighting of effect size estimates based on sample size was necessary. To do this, all effect sizes originally calculated as Cohen's *d* were corrected for sample size into Hedge's *g* so that effects in small samples were not overestimated.

Although a full vote-count method (Hunter & Schmidt, 2004) was not employed, the direction for each effect was reviewed to identify overall trends. Effect sizes that were positive indicated that the treatment group scored better than the control group whereas a negative effect meant the comparison group outperformed the treatment group. In the cases of 11 of the effect sizes (17%), the control group outperformed the treatment group. This was consistent with other reviews which found negative effects in 16% of studies reviewed (Lehr et al, 2003). The Pearson product-moment correlation between year of publication and *g* was $-.398$ ($df=64$, $p < .01$) indicating that there was a significant negative correlation between these two variables; namely that more recent studies tended to report smaller magnitudes of treatment effects. Although this result has several possible explanations, it has been suggested in other reviews (Agodini & Dynarski, 2004) that historical publication bias against non-significant or weaker findings relating to

dropout prevention interventions might be lessening. Additional analysis revealed however, that no journal article reported an effect size higher than ± 0.47 and although technical reports and dissertations both generally reported larger magnitude of effects, dissertations in particular reported large effects and none of the dissertations reported negative treatment effects.

Effect sizes extracted for Canadian-based programs as compared to Non-Canadian programs revealed (see Appendix F) that Non-Canadian based programs focused more frequently on measuring program impacts in terms of “physical presence” (e.g., attendance, dropout, graduation) whereas Canadian programs had a more diverse approach to measuring program impacts, including social/behavioural (e.g., problem behaviour, social competence) and psychological outcomes (attitudes towards school, self-esteem)

The interventions analyzed in this review varied widely in constructs. No two reports studied the same particular intervention (although Hayward (1995) reported on multiple sites with the same intervention model) and even among studies categorized under the conceptual outcome category of “physical presence”, some studies reported on graduation rates, others absence rates over a semester, others enrolment status 1, 2 or even 3 years after the intervention. As Lipsey and Wilson (2000) point out, because meta-analysis techniques focus on the aggregation and comparison of findings “it is necessary that those findings be of a sort that can be meaningfully compared....that is, deal with the same constructs and relationships” (p. 2). Thus, although the conceptual construct of “physical presence” was used as a way to group related findings, there may be limitations to operationalizing an overarching concept when it comes to aggregating the findings.

In addition, since several reports provided multiple effect sizes or, as in the case of Mutadi (1990) where the same control group was compared to multiple treatments groups, there was an issue in aggregating effect sizes as to whether the assumption of statistical independence among the data points could be maintained. Additionally, where one report, Hayward (1995), accounted for 33% of the effect sizes extracted (n=22), there was concern that the one report may give more weight to the overall analysis than the others. For these reasons analysis was undertaken to explore the assumption of homogeneity and the influence of outliers.

The weighted effect sizes were aggregated to form an overall weighted mean estimate of the treatment effects on the subsets of outcomes (i.e., g+). The significance of the mean effect size was judged based on its 95% confidence interval (i.e., does the interval contain a value that implies no effect?) and on whether the CI was partly or entirely within what could be called a range of clinical indifference (Ogles, Lambert, Masters, 1996), that is, does the magnitude of change produced by the treatment fall within a range that does not suggest existing practices be changed. Although there is growing number of researchers who advocate using the random-effects model (e.g., Hunter & Schmidt) in meta-analysis, a fixed model was adopted in this case because the random effects assumption fundamentally asserts that the analyst addresses the question, “will the treatment produce benefit ‘on average’”, whereas the fixed effects assumption leads to addressing the question “did the treatment produce benefit on average in the studies at hand?” (Petitti, 2000). In this review the interest was not in whether dropout prevention programs *will have* an effect but whether the treatments being reviewed *have*

caused an effect. Table 8 shows the results for all 66 effect sizes grouped by outcome category.

Table 8. Weighted Mean Effect Sizes for Combined Outcome Categories

Outcome Category	Effect size		95% confidence interval		Homogeneity of effect size	
	g+	SE	Lower	Upper	Q value	df
Physical Presence	.067	.018	.031	.103	230.374	48*
Academic/Cognitive	.209	.052	.107	.310	21.558	3*
Psychological	.495	.052	.394	.596	64.812	8*
Social/ Behavioural	-.041	.083	-.203	.121	4.260	3**

**p < .05

** P < .5

Overall, three outcome categories showed positive and expected treatment effects, with a modest positive net effect size of .495 for the small number of psychological outcomes. Further examination of psychological outcomes revealed that several of the outcomes fell outside + 3.0 standard deviations from the weighted mean effect size for psychological outcomes and could have been considered outliers. For example, one effect size from the Mutadi study was more than double the magnitude of the next largest effect size extracted among psychological outcomes. There were no effect sizes for psychological outcomes demonstrating a negative treatment effect.

Sampling error was initially suspected in this case due to the fact that extreme values can occur simply because of large sampling error from small sample sizes. Although the trend in behavioural and social sciences has recently been to eliminate up to the most extreme 10% of data points (the largest 5% and the smallest 5% of values)

(Hunter & Schmidt, 2004) to reduce the impact of outliers, that approach is not always feasible. In cases such as this, where you are working with small to moderate sample sizes, these extreme values can simply occur due to sampling error and might not be “true” outliers (Hunter & Schmidt, 2004). It is because of the problematic nature of working with potential outliers within small subsets of data that caution should be taken in interpreting the overall effect of interventions on this set of psychological outcomes.

The weighted mean effect size for the 49 physical presence outcomes, the measure perhaps most relevant to exploring program effectiveness in this case, was essentially zero (.067), but the test of homogeneity ($Q=230.374$) indicates a certain degree of variability surrounds it. That is, the actual effect size of the population could range substantially on either side. The range of effects sizes for physical presence outcomes (-0.656 to +1.053) suggests that some interventions have statistically been more successful than others. That said, researchers in education continue to discuss how to determine the practical significance of an effect size calculation. Cohen (1988) recommended that $d=.20$ (small effect); $d=.50$ (moderate effect), and $d=.80$ (large effect) as a general guidelines across a variety of disciplines. Many educational researchers (Gall, Borg & Gall, 1996) consider an effect size of .33 as the minimum to establish practical significance. Using Cohen’s guidelines as a benchmark example, we see that 77% of the total individual effects were less than $d=.50$ (moderate) and 36% ($n=24$) were less than $d=.20$ (small). Mutadi (1990) presented a unique case in that all 6 of the effect sizes extracted were $d=.72$ or higher.

In exploring additional features of these studies that might provide a better understanding of the range of effect sizes observed, only the outcomes coded as “Physical

Presence” were analyzed further because the number of cases for all other outcome categories were extremely small (Hunter & Schmidt, 2004).

A growing trend in dropout prevention interventions is incorporating multiple instructional components in a single program that address not only academic needs but health, life skills, and cultural/spiritual enrichment (Lehr et al., 2003). To explore the issue of whether more intense and complex interventions reviewed here are different from more basic and simple forms of intervention a comparison of overall mean effect sizes was compared for findings involving 3 program components or less (i.e., simple) and those incorporating 4 or more (i.e. complex). Table 9 provides a summary of the analysis. Although comparisons of weighted mean effects (i.e., g values) between the two groups suggests that programs incorporating multiple components are demonstrating stronger effects, the between group analysis was not significant ($p=.219$), therefore, there was no statistical evidence that multi-component interventions have had any greater effect on physical presence outcomes.

Table 9. Weighted Mean Effect Sizes Based on Program Complexity

Level of Complexity	Effect size		95% confidence interval		Homogeneity of effect size	
	g+	SE	Lower	Upper	Q value	df
Simple	.052	.022	.008	.095	132.969	27*
Complex	.101	.033	.036	.165	95.893	20*

*p < 0.000

Additionally, the sample of effects sizes measuring physical presence was split into Canadian and Non-Canadian based programs to explore potential differences. Results indicated that the average mean effect on physical presence outcomes was significantly higher ($p=.003$) for Canadian programs. See Table 10 for a summary of results. However, it should be noted that the number of cases used in calculating the Canadian average was quite small and had significant heterogeneity ($p<.0000$), thus caution should be exercised in interpreting these findings because of the low number of outcomes associated with Canadian-based programs.

Table 10 Weighted Mean Effect Sizes on Physical Presence Outcomes for Canadian and Non-Canadian Programs

Country	Effect size		95% confidence interval		Homogeneity of effect size	
	g+	SE	Lower	Upper	Q value	df
Canada	.247	.062	.125	.369	65.404	9*
Outside of Canada	.050	.019	.012	.088	155.859	38**

*p < 0.000 ** p< 0.010

In summary, overall analysis of effect sizes was limited by the small number of reports including numeric data on program impacts and on the complexity of outcome categories. The most important outcome of the overall analysis of effect sizes was the wide variability both between the outcomes being measured and within subsets of

outcomes. While the average effect size for outcomes related to students' physical presence was near zero, there was a range of effects reported, from moderately negative to strongly positive.

The persistent heterogeneity of measures indicates that taking these findings as a true representation of the entire population values may be risky. It would be irresponsible to suggest that dropout prevention programs have had a significant impact on physical presence if you are to look solely at mean effect size and heterogeneity. When faced with such wide and unexplained variability, as seen in many other reviews (e.g., What Works Clearinghouse, 2008) it is unfortunate that more studies did not report impact data so that a more comprehensive examination of program impacts could have been undertaken.

It should be noted that even among the non-empirical studies included in this review the quality of descriptions on specific outcomes and program impacts was often far inferior to the quality of descriptions related to implementation challenges and discussions of dropout correlates. The dearth of descriptive and empirical outcome related data resulted in having to recode several of the study features because there was not enough variability among studies to make those categorizations practically meaningful. For example, coding the extent to which particular curricular strategies were employed in relation to others (e.g., Item # 33 service learning, Item #32 cooperative learning, tutoring, Item # 34 mentoring, etc) were simplified from a "more/equal/less than" model of coding to a simpler categorization of "yes/no" to aid in analysis.

While the analysis of effect sizes presented above highlighted some interesting results that can inform current understandings, it was only one component of the review and not meant to be taken as the complete analysis of studies under review. The

following sections provide a more global overview and analysis of the data collected for this review, including results and findings from the full range of research design (i.e., both empirical and case-study designs). It was hoped that by exploring general trends and patterns beyond effect size analysis that the research objectives relating to increasing awareness and common understandings about current practices might be better achieved.

Demographic Features

RESEARCH DESIGN

About 68% (n=103) of the total 151 programs and interventions were broadly defined as using a “case study” research design. These case study designs varied from evaluation studies using mixed methods of analysis to qualitative single site case studies applying a more action research or even narrative approach. As evidenced in Table 11 reporting on Canadian based programs was more frequently identified as case study design than those reviewed from outside of Canada. Studies that incorporated random selection and assignment of participants to groups or established comparison groups without random assignment made up about 28% (n= 42) of the total group of studies reviewed (Canadian and Non-Canadian combined). However, it should be noted that very few Canadian findings were identified as employing random or non-random assignment.

Table 11. Research Designs Compared Between Canadian and Non-Canadian Based Programs

Research Design	Canadian (n=95)		Outside of Canada (n=56)	
	Frequency	%	Frequency	%
Random	5	5%	24	43%
Non-random	7	7%	6	11%
Pre-post	1	1%	5	9%

Case-study ¹	82	86%	21	37%
-------------------------	----	-----	----	-----

¹ Case study was recoded to broadly include any in-depth study of an intervention or program that was mainly descriptive in nature such as action research, ethnographic, or evaluation studies which focused on the experiences of participants
N=151

SELECTION CRITERIA FOR PROGRAM/INTERVENTION

One hundred and twenty-three of the total group of findings reported participant selection criteria (i.e., 82 %), a significant number of the programs selected participants based on multiple criteria. The most common criteria for eligibility was history of academic performance; 42% (n=63) reported GPA, test scores or some other measure of academic performance as the main selection criteria. Next, teacher referral was utilized in 11% of the findings (n=16). The remaining selection criteria were all used less than 10% of the time; history of dropping out of school 9% (n=13), behavioural or psychological issues 6% (n=9), attendance records 6% (n=9, age 4% (n=6) and socio-economic background 5% (n=7). Table 12 provides a breakdown of selection criteria cited in Canadian and American programs. Selection criteria followed similar patterns among the Canadian and American programs, with a slightly higher percentage of Canadian programs using teacher referral as the point of reference for student selection.

Table 12. Selection Criteria for Canadian and Non-Canadian Based Programs

Main Selection Criteria	Canadian (n=95)		Outside of Canada (n=56)	
	Frequency	%	Frequency	%
Academic performance	33	35%	30	54%
Attendance records	5	5%	4	7%
Teacher referral	15	16%	1	2%
History of dropout	7	7%	6	11%
Socio-economic status	2	2%	5	9%

Social/behavioural issues	7	7%	2	3%
Criteria not reported	20	21%	8	14%

PARTICIPANT CHARACTERISTICS

Less than half of the total number of findings (n= 63) reported on the geographic location of the program or intervention. Of those interventions, a slight majority 57% (n= 36) took place in urban settings with the remaining 43% (n=27) taking place in a rural communities. Only a few cases (n=32) described or reported on the ethnicity of participants; aboriginal and Hispanic youth were the only two ethnic groups specifically identified. Some reports described participants as “ESL students”, “immigrants” or “racial minorities” but did not provide further detail. Gender, SES, age and other individual participant related variables were virtually unreported across studies.

Program Features

SCHOOL LEVEL

Almost all of the cases, 94% (n=142), reported on the school/grade level of the participants. When examining the combined data from Canadian and American programs, the majority were targeting students at the secondary level 44% (n=66) or junior high and senior high level combined (25%, n=37). Twenty-four percent of the intervention findings reflected programs geared towards junior high school students alone. Only two percent of the overall findings (n=3) emerged from programs at the elementary or preschool level; all of which were Canadian-based programs. It is worth noting in Table 13, that Canadian based programs were more often identified as targeting students in

Junior High and younger, whereas American based programs tended to focus more on Junior High and older.

Table 13. Comparison of Targeted Grade Level

Grade Level	Canadian (n=95)		Outside of Canada (n=56)	
	Frequency	%	Frequency	%
Pre-School	1	1%	0	0%
Elementary School	2	2%	0	0%
Junior High School	28	29%	9	16%
Junior/Senior Mixed	25	26%	11	20%
High School	30	32%	36	64%
Not reported	9	10%	0	0%

PROGRAM TYPE

After coding was completed, results indicated that the most frequent model of program delivery was to pull students out of their main classes for specialized support (39%, n=59). Least reported were interventions designed to be used by regular classroom teachers in class (5%, n=7) and entirely separate cooperative education programs (3%, n=5). Table 14 summarizes differences observed in program types for the total group of findings (n=151).

Table 14. Reported Model and Schedule Designs for All Programs

Program Type/Model	No.	%
Pull-out programs	59	39%
School within a school	18	12%
Alternative school	29	19%
Workshop or special course	25	17%
Cooperative education	5	3%
In class intervention	7	5%
Insufficient description to categorize	8	5%
<hr/>		
Program Schedule		
Regular Set Time	54	36%
Flexible hours	13	8%
Drop-in centre	1	1%
Individualized	20	13%
Did not report	63	42%

It is worth noting that Canadian based programs presented a more varied approach to program types (see Table 15) with a higher percentage of Canadian programs being delivered in alternative schools, through workshops or special classes or through changes within the existing classroom. Descriptions of how the intervention or program was scheduled were insufficient for coding in only 42% of the cases (n=63). Among cases where that information was reported 61% (n=54) involved programs that were delivered using a regular set schedule. Fifteen percent (n=13) followed a flexible timetable for program delivery and 23% (n=20) incorporated individualized timetables (See Appendix B for descriptions of each schedule code).

Table 15. Comparison of Program Models for Canadian and Non-Canadian Based-Programs

Program Type	Canadian (n=95)		Outside of Canada (n=56)	
	Freq.	%	Freq.	%
Pull-out program	38	40%	21	38%
School within a school	3	3%	15	27%
Alternative school	21	22%	8	14%
Workshop/class	19	20%	6	10%
Cooperative education	4	4%	1	2%
In class	7	8%	0	0%
Insufficient description	3	3%	5	9%

The structure for scheduling the interventions and related issues of student-control over the learning process was one of the only themes emerging from the secondary qualitative analysis in this review. Initial qualitative results demonstrated that the degree to which programs individualized support for students was a contentious issue. Participants and stakeholders in many of the programs reviewed felt strongly that individualized instruction “facilitated self-directedness”, “allowed an equal and fair share in decision-making” and that “self-imposed deadlines with guaranteed individual help with no fear of reproach” and “free choice [were] the most important factor underlying their success”. By not allowing for individualized instruction many stakeholders felt programs were “replicating some of the mainstream schooling practices that had alienated students”. Qualitative coding and detailed notes revealed, however, that many stakeholders across a variety of these programs were concerned with the degree of student choice and self-direction incorporated into the instructional design, saying; “students don’t take it seriously” and that self-directed programs “need a lot of work to provide a challenging program to students” and that standardization was necessary

because “mediocrity won’t allow them to survive in the real world”. Although attempts were made to explore demographic and program features within qualitative findings that might further explain these differing viewpoints the lack of descriptive narratives and data made this very problematic.

DURATION OF INTERVENTION

The duration of an intervention, comparable to the dosage measure of a treatment in medical research, is important in education. A measure of duration was included in this review to identify the amount of time students were provided support. In the large majority of instances (83%) the interventions or program involved exposing students to support/treatment lasting more than 1 semester or 12 weeks in length (n=125). However, program design varied significantly with some programs engaging students daily over the course of 12 weeks or more and other programs engaging them once a week or even less frequently over the course of the treatment timeframe. The level and intensity of program intervention is not necessarily reflected in program duration. A few studies 3% (n=5) involved interventions that were less than 12 weeks long. These were normally workshops, summer adventure-type programs, or specialized intense training over several days. Only 51% of cases reported the number of years the intervention had been running before the data being reported was collected. Of the cases which did report this information, 83% had been in operation for 3 years or less at the time the results were published.

PROGRAM COMPONENTS

After coding was completed results indicated the major component (80%, n=120) of most programs generally was on providing academic support. See Table 17 for summary of

overall findings and Table 16 for breakdown between Canadian and American programs. Life skills (e.g., personal care, behaviour modification, budgeting, time management) 62% (n=93) and vocational training 42% (n=64) were also strong focuses of programs reviewed. While no programs incorporated all seven program components, seven (5%) incorporated almost all of the program components coded for. The majority of programs reviewed did, however, deliver educational programming across multiple dimensions. For example, thirty-two percent of cases (n=49) were coded as incorporating 4 or more program design feature (i.e., health, life-skills, career prep, etc) and 74% (n=112) were identified as incorporating at least two

Table 16. Program Components for Canadian and Non-Canadian Based Programs

Program Component	Canadian (n=95)		Outside of Canada (n=56)	
	Frequency	%	Frequency	%
Health	24	25%	20	36%
Life-Skills	58	61%	35	63%
Career Preparation	36	38%	20	36%
Academic Support	65	68%	55	98%
Cultural/Spiritual	10	11%	15	27%
Parental	16	17%	17	30%
Vocational Training	33	35%	31	55%

* Note: percentages do not equal to 100% due to rounding

Comparisons between Canadian based programs and those outside of Canada revealed similar patterns of emphasis on program design components with academic and life skills being most frequently reported by both groups. Vocational training and academic support were more frequently noted as components of programs outside of Canada.

INSTRUCTIONAL DESIGN

As shown in Table 17, the nature of instructional activities and strategies varied across findings with the most common instructional strategies being tutoring (63%, n=95), work-based learning (50%, n=75) and mentoring (40%, n=61). Least reported instructional strategies were service learning (7%, n=10), multiple forms of assessment (9%, n=13) and culturally responsive teaching (12%, n=18). Just over half of the findings from programs targeted towards Aboriginal students reported using culturally responsive teaching strategies (57%, n=8). Lack of detailed reporting of program features was especially disappointing as features of how programs were designed and the kinds of teaching practices and student work involved are a critical component of furthering out understanding of best practices.

Table 17. Summary of Program Components and Instructional Strategies Across All Interventions

Program Components	No.	%
Health	44	29%
Life-Skills	93	62%
Career Preparation	56	37%
Academic Support	120	80%
Cultural/Spiritual Enrichment	25	17%
Parental Development	33	22%
Vocational Training	64	42%
Instructional Strategies		
Culturally Responsive Teaching	18	12%
Self-expression	39	26%
Multiple forms of assessment	13	9%
Cooperative learning/group work	31	21%
Service learning	10	7%
Mentoring	61	40%
Tutoring	95	63%
Computer Supported Learning	29	19%
Work-Based learning	75	50%
* based on entire sample of Canadian and Non-Canadian based programs n=151		

Table 18 summarizes the instructional strategies individually within Canadian and American programs. Although instructional programs seemed generally to follow similar patterns, Canadian based programs more frequently reported instructional strategies involving multiple forms of assessment (14%), and self-expression (32%) whereas American programs more frequently reported incorporating computer supported learning (36%), work-based learning (64%) and a stronger emphasis on tutoring (82%) within the instructional activities.

Table 18. Instructional Strategies Used by Canadian and Non-Canadian Based Programs

Instructional Strategy	Canadian (n=95)		Outside of Canada (n=56)	
	Freq.	%	Freq.	%
Culturally Responsive	9	10%	9	16%
Self-expression	30	32%	9	16%
Multiple forms of assessment	13	14%	0	0%
Cooperative learning/group work	21	22%	10	18%
Service learning	3	3%	7	13%
Mentoring	37	39%	24	43%
Tutoring	49	52%	46	82%
Computer Supported Learning	9	10%	20	36%
Work-Based learning	39	41%	36	64%

SUPPORT STAFF

The level and nature of instructional support available to students took a variety of forms. Results of all studies (n=151) generally indicated that outside of the teacher Psychologist/Counsellors (31%, n=472), Mentors (29%, n=44), and Peers (19%, n=29) were the most common forms of additional staff support. To a lesser extent Social Workers (15%, n=23), Substance Abuse Counsellors (8%, n=12), Parents, (17%, n=25) and Health Care Practitioners (13%, n=19) were involved in program delivery. Most program findings (70%, n=106) also reported the involvement of additional staff, although the kinds of staff support varied significantly across studies, for example: youth care workers, teaching assistants, business professionals, or specific administrative support personnel, were cited as required support staff. That said, while the availability of those support personnel was reported the extent to which learners accessed those support systems or the quality of those interactions was not reported in any depth.

IMPLEMENTATION CONSIDERATIONS

Secondary analysis of qualitative notes revealed a key theme clustered around factors identified by stakeholders as affecting successful implementation of instructional strategies. A recurring theme across the full range of findings was an underlying emphasis within the program design on creating a positive *psychological* experience for learners. Although many studies did not state this as an explicit program objective deeper qualitative analysis revealed that stakeholders were challenged between seeing the value of dropout prevention programs in helping learners develop confidence and self-esteem and seeing the programs as “devaluing” students by taking them outside the “mainstream” formal education process. The following table outlines a sample of the kinds of participant descriptions that were given relating to underlying psychological values and barriers in implementing dropout prevention programs.

Table 19. Qualitative Descriptors of Program Value and Barriers

Sample of descriptions of the psychological value of dropout prevention programs
<ul style="list-style-type: none"> • Teachers feel “they are more successful at improving self esteem then engaging them academically” • Students “need a lot of encouragement and self-esteem building” • [program] promotes self esteem. • [program] promotes self-awareness and self-assertion through social interaction • Self esteem ratings of students rise when they drop out of school...programs like this help • Gives them back some of the pride they lost by not being able to continue in their regular school • [program supports] the maintenance of those human characteristics that foster approval and success • Disconnections between thinking and feeling can be addressed • [program] foster respect for aboriginal culture and traditions • [programs] help us to stop blaming the victim for needing assistance
Sample of descriptions of psychological barriers for dropout prevention program
<ul style="list-style-type: none"> • “Located on the margins of conventional schools” • Devalued • Label of “rehabilitation” program • Implies that “students are broken and need to be fixed” • Remedial curriculum stigmatizes the program as second class • The more {the program} aligns itself with the competitive academic curriculum then it challenges the stigma of being a lower stream • {programs} aim to ‘rehabilitate’ these ‘problem youth’ through the eradication of their assumed deficiencies often paying minimal attention to economic and social conditions • Program could improve by not being ‘fake’ • There is a stigma attached to attending the program for “slow learners” • Should not been seen as a ‘dumping ground’ and referral streaming needs to be improved • Programs continue to struggle against a societal devaluing of non-white-collar careers

Program Effectiveness

Dependent variables were determined by examining what measures were used to assess levels of success or failure. Although all studies were selected because they were self-described as a dropout prevention program, not all studies measured outcomes addressing enrolment status. Although effect sizes were extracted for a number of studies many reports provided narrative and descriptive results. To be able to explore general trends across all studies in terms of program outcomes, including the case studies, both the type of outcome measured and the strength of evidence reported for each of the 151 program interventions reviewed were coded. Although “extent of evidence” guidelines, used by organizations such as the What Works Clearinghouse, are valuable and vetted protocols, those guidelines deal specifically with reviewing empirical studies. The purpose of this review was to also include findings outside of the published empirical literature so instead a coding scheme which attempted to reflect on key factors to consider when evaluating the strength of program evidence drawing on recommendations by Smink (1992) was conducted. This was more than a measure of validity (which is discussed later) but a way to categorize the outcome evidence reported across all studies. The coding scheme is detailed in the codebook, Appendix B.

OUTCOME CATEGORIES

Because the range of outcomes was so diverse, five broad categories based on the model presented in Lehr et al. (2003) were used to cluster the indicators of program effectiveness: Academic/cognitive = GPA, standardized test scores, study habits, etc., Physical presence = enrolment status, attendance, graduation rate, dropout rate, etc.,

Psychological = student attitudes towards learning or school, self-esteem, depression, etc., Social Behavioural = problem behaviour, social competence, drug use, violent behaviour, etc., and Support for learning = student attitude toward teachers, school climate. The majority of findings primarily measured changes in students being physically present at school (70%, n=105). About 17% (n=26) of the interventions primarily focused on reducing dropout by improving psychological outcomes like attitudes toward learning and school, self-esteem, depression, etc. Only 3 % of the programs had outcome variable pertaining to social/behavioural (n=5) or support for learning effects (n=4).

Table 20. Categories of Program Objectives for Total Sample (n=151)

<i>Primary Outcome Sought</i>	<i>No.</i>	<i>%</i>
Academic/cognitive	11	7%
Physical presence	105	70%
Psychological	26	17%
Social/behavioural	5	3%
Support for learning	4	3%

STRENGTH OF OUTCOME EVIDENCE

Keeping in mind that in general the validity of studies was quite low, interventions were coded for the extent to which the authors presented evidence of program effectiveness (i.e. strong positive, positive, mixed results, negative, or strong negative results). Global analysis of reported program evidence suggested that almost three quarters of interventions reported positive program effects (72%, n=109) with another 13% (n=19) reporting *strong* evidence of positive program effectiveness without any

overriding contrary evidence. Several of the programs reported inconsistent or mixed effects (10%, n=16) and seven interventions demonstrated only negative effects (5%) suggesting the program was ineffective.

Table 21. Summary of Effectiveness Evidence for Total Sample (n=151)

<i>Strength Of Program Effectiveness</i>	<i>No.</i>	<i>%</i>
Strong Positive	19	13%
Positive	109	72%
Mixed/Inconsistent	16	11%
Negative	7	4%
Strong Negative	0	0%

The studies reporting mixed effects often in trying to triangulate data reported different kinds of results emerging from different data collection techniques. For example one study by Campbell (1995) made the following observation:

“inconsistencies in the evaluation results of the breakaway company described thus far raise questions that beg answers. Why do these students give answers to the scaled self-reports measures that do not appear to match or support what they say about themselves in an open interview format? What are the scaled measures and interviews assessing, if not the same things? And, what is the relationship between the timing of observations of these students’ behaviours and the quality of their behaviour? When are observations trustworthy for evaluating program outcomes?” (p. 448).

Further analysis of outcomes revealed that of the 19 interventions demonstrating strong positive effects, 90% incorporated tutoring as an instructional strategy (n=17), 58% involved career and technical forms of instruction (n=11) and 47% integrated a mentoring as part of instructional delivery (n=9). In similar fashion, the seven

interventions demonstrating negative effects were of similar composition, 86% (n=6) included a tutoring component, 71% (n=5) a vocational component, and 43% (n=3) included a mentoring component. Table 22 displays the results for instructional strategies implemented. In addition to instructional design strategies, location and program components did not differ significantly between the groups of negative and high performing cases.

Table 22. Summary of Instructional Strategies by Outcome Strength

Instructional Strategy	Strong Positive Outcomes (n=19)		Negative Outcomes (n=7)	
	Freq.	%	Freq.	%
Culturally Responsive	2	11%	2	29%
Self-expression	2	11%	2	29%
Multiple forms of assessment	0	0%	0	0%
Cooperative learning	3	16%	2	29%
Service learning	5	26%	0	0%
Mentoring	9	47%	3	43%
Tutoring	17	90%	6	86%
Computer Supported Learning	6	32%	1	14%
Work-Based learning	11	58%	5	71%

Results comparing general methodological validity and general strength of outcomes also revealed interesting patterns. As seen in Table 23 in all seven cases where only negative results were reported, the study demonstrated high internal validity. However, in the case of inconsistent or mixed results, nearly all of them were from studies of lower research validity.

Table 23. Summary of Validity Ratings Compared by Outcome Strength

<i>Strength Of Program</i>	Validity			
	High (n=43)		Low (n=108)	
	No.	%	No.	%
Positive	35	81%	93	86%
Mixed/Inconsistent	1	2%	15	14%
Negative	7	16%	0	0%

Summary and Discussion

The purpose of this study was to provide a descriptive overview of the current research, especially within the Canadian context, relating to instructional interventions to reduce dropout, to evaluate those programs, if possible, for underlying factors that may help explain program success or failure, and to increase awareness about implementation issues that affect how stakeholders might design and implement programs in the future. The following discussion aims to highlight the contribution this review has made and to summarize what information has been gleaned.

How are Programs Measuring “Effectiveness”

It has been firmly established in the research literature that there are many paths and factors which lead a student to drop out of school, some of which are instructionally related, many of which are not. The present review confirmed that instructional interventions within Canada, specifically, and abroad more generally, are being implemented with a variety of focus. While students physical presence in school (i.e., absence rates, dropout rates, graduation rates) is one way that programs attempt to

measure success, other programs look at assessing improvement in psychological barriers to school completion (e.g., attitudes toward school, self-esteem) or improvements in dropout correlates (e.g., negative social behaviour).

This diverse focus was particularly true of Canadian-based programs reviewed in this study, with Canadian based program more frequently measuring social behavioural and psychological outcomes than their non-Canadian counterparts. Although many of the Canadian studies reflect current thinking on important constructs (student engagement, attitudes towards self and schooling, etc.) there is a need to link them more directly to impact data. How “effectiveness” and “good performance” is defined and eventually measured by stakeholders of interventions affects not only our ability to generalize about program success but also to explore best practices. If a program designed to reduce dropout rates focuses on improving social behavioural characteristics of participants then “what works” might be very different than “what works” for a program designed to reduce dropout rates by improving core academic skills. Although all interventions included in this review were designed as dropout prevention or intervention programs, thirty percent of studies did not collect or report on outcome data related to students presence at school at all. Moreover, even among programs categorized as measuring physical presence, many did not report dropout or graduation outcomes specifically but on related constructs like attendance rates. Frequently, percentage of dropout was provided as a measure of success without any explanation of how those “dropout rates” were computed or discussion of the benefits/limitations of using that particular approach.

Only 48 effects sizes related to students physical presence at school could be extracted and half of those were reported in a single study by Hayward (1995) who was

assessing a range of vocational programs based outside of Canada. If the purpose of a program is to prevent students from leaving school, outcomes should include actual changes in dropout or graduation rates, even if other dependent variables are relevant. Though measures of student engagement, attitude and indicators of effectiveness relating to dropout predictors are valuable in developing our understandings of variables that affect at risk learners, measures of school completion and even post-school outcomes (i.e., post secondary attendance, work outcomes) are fundamental for making conclusive statements about the effect of these instructional programs on learner performance and success.

Whether dropout rate data was collected but just not reported due to political or bureaucratic strategizing is obviously not known, but the lack of published and publically available or even identifiable studies investigating program effectiveness in terms of dropout out and graduation rates specifically is particularly concerning when currently in Canada all levels of government as well as the private sector are making claims about substantial funding initiatives being targeted at interventions aimed at increasing school completion (e.g., Ontario Lighthouse Project). The utilization of evaluation results is the final test of the quality of the evaluations being conducted across the country. The usability of findings is directly related to the credibility and accessibility of the evaluation results and the relevance of the data to support improved practice. Teachers, parents, and community leaders need to know if these programs are having an impact on overall school dropout rates, the business community needs to know if cooperative education and vocational opportunities for students are in fact helping students stay in school and develop marketable work skills. Funding agencies need to know whether the components

of programs they are financing are contributing in a valuable to overall program results (Smink, 1992). Most importantly the true stakeholders in these programs, namely the students, need to know that they can benefit from the interventions being provided. The responsibility of evaluators and policy-makers in this regard cannot be underestimated.

As Smink (1992) noted,

“it is critical to the future of dropout prevention programs that policymaking and funding agencies receive information about the effectiveness of various strategies in reducing the student dropout rate... the data needed for this kind of informed decision-making would be available is all dropout prevention programs included an evaluation process in their program planning, p.233)

Lack of impact and outcome data is not exclusive to Canada however. Other reviewers have observed similar problems (e.g., Lehr, 2003), arguing that despite increased pressure within the educational system to be results-oriented even the published reports are insufficient in their descriptions of program design/implementation and lacking in their analysis of program impacts to conduct meta-analytic or additional kinds of analysis to statistically demonstrate “what works”. Even 20 years ago, Bickel (1991) concluded that “a central weakness ... (is) the lack of solid evidence about what is being accomplished by the programs(s). We are strong on description and testimony and short on hard evidence” (p. 74).

Although there is a vast literature discussing the many complexities and the importance of instructional interventions for children and youth at risk of dropping out the majority appear to be conceptual in nature. It appears that in the last decade and a half since the last major review conducted in Canada by Morris, Pawlovich & McCall, D (1990) the repeated calls for improved reporting on educational initiatives relating to dropout prevention interventions in Canada continues to go largely unheeded (Jordon,

2006, Lehr et al, 2003; Prevatt, 2003, etc.). Further to that point, evaluations should not be built around rationalizations of why a program did not meet expected outcomes. Assessment of situational and environmental factors should be expected in advance of the intervention. Findings from this review reaffirm the need for creating avenues for sharing and discussing existing programs and for improving the methodological quality of program evaluations being conducted.

Another related contribution of this review is supporting the notion that common measurement tools which can be applied across program contexts to assess outcomes are badly needed in this literature. In a recent report “Who’s Counted, Who’s Counting: Understanding High School Graduation Rates” (Pinkus, 2006) the authors contend that there are a staggering number of calculation formulas being used both in Canada and in the US for determining graduation rates and that this is negatively affecting the way researchers and analysts determine the “effectiveness” of individual intervention, but also how they calculate the overall provincial/state-wide, and federal dropout rates. In this review, there were a range of strategies used to assess dropout rates; for example: a) Hayward (1995) used cumulative dropout rates (i.e., assessing the number of dropouts in the treatment group at the end of the third year of the intervention); b) Hahn (1994) used self-reported status 4 years post-intervention; and c) McPartland (1991) calculated absence rates for the second year of the intervention by dividing the number of days absent by the number of days each student was on the official roll call. Moreover, often times authors referred simply to “dropout rates” but did not provided detail on how those rates were determined (e.g., using board/district level enrolment data, soliciting self-

report information from former participants, designating an individual with a certain percentage/number of consecutive absences as a “dropout”) .

Not only does this lack of consensus about how to measure dropout obscure the graduation rate crisis in many communities, particularly for at-risk students and minorities, but it also makes comparisons across provinces, programs, schools, etc. nearly impossible. Having a standard method for calculating and measuring dropout rates is a critical factor for policy makers, designers and other stakeholders in figuring out where to allocate resources and develop support services. Lack of consistency in measuring program effects raises doubt about the efficacy and effectiveness of programs, despite the fact that many programs in this review are claiming success. As Lehr (2003) noted; “grouping interventions and dependent variables using a consistent framework may assist in understanding intervention effectiveness” (p. 360). This review reinforces recent calls for a more standardized approach to dropout calculation and for those formulas and approaches to be publically reported to ensure comparable measurement methodologies are being used.

Methodological Quality of Current Research

Evidence-based practice is more than just an accumulation of professional anecdotal knowledge; it should incorporate empirical evidence as well. This review found that case-study designs were the predominant methodology used in evaluating and reporting on programs (i.e., 76% of Canadian programs) and that empirical design were infrequently employed. Although the practical and ethical difficulties of using randomized research designs with at-risk populations is obvious, it does bring into

question how large a sample size is needed to judge the result of a program as credible. Many of the programs reviewed in this study reported small sample sizes (e.g., less than 30 for the treatment group) which often was a reflection of a school-based initiative or local program developed independently of other programs across the country.

When at-risk populations are targeted, it may be difficult to have large sample sizes for testing, however, if the goal is to test the efficacy of an intervention, sample size should be considered in advance. Detailed reporting of impact data need not be a by-product of research design choice. Even in the case of comparison group designs more could have been done to limit the potential for erroneous conclusions based on poor research design. Very few comparison group designs established equivalency via non random design or conducted any kind of post hoc analysis for group variability. In the future researchers should, in these cases, collect data which assesses group comparability on things like geographic location, socioeconomic and other personal characteristics, time period in which they were studied, and the methods used to collect the data (Baron, 2002). This review suggests that evaluators and designers need to carefully consider the process for how individual students are identified for intervention or prevention support and how they might capture some of the unobserved factors (e.g., motivation) that may influence outcomes.

Even in cases of random assignment (e.g., Breckon, 1996) where selection criteria were used to identify high risk students, then students were randomly assigned to either control and treatment groups; virtually no descriptive data was provided detailing individual and group characteristics (i.e., gender, academic standing, SES) that might help in interpreting and understanding potentially pre-existing differences within groups.

While some researchers caution strongly against using well-matched comparison group designs “as a final arbiter of what is effective and what is not” (Baron, 2002, p.12), they may provide the most realistic and accessible option for researchers, particularly when dealing with the ethics involved in random-assignment of at-risk learners.

Qualitative researchers in this field need to work at developing commonly defined constructs and methods for exploring program outcomes as well. To date qualitative research has been underused as an evidence resource relating to evaluating dropout prevention programs and some argue this is mainly due to the underdeveloped science and methods applied in conducting that kind of analysis. This was certainly the case in the present review. Although a large majority of studies reviewed focused on providing descriptive analysis of programs, primary studies did not provide sufficient theoretical and methodological description to allow for more than a cursory analysis of methodology and research findings.

The five point likert scale used in this review to assess global measures of outcomes, including case study designs, was useful in that it allowed for a more inclusive discussion of both quantitative and qualitative outcomes and conceptualizing of the broader data set. However, because there was such variety in terms of design, internal validity and types of evidence used to determine the strength of a reported outcome you ended up grouping results that were measured with results that were perceived (either by participants themselves or by the authors). The proverbial “apples and oranges” debate in systematic reviews will undoubtedly continue but in this instance, where the review was meant to explore the range of impacts and treatments and to engage the larger community involved in this issue, finding a way to synthesize both quantitative and qualitative studies was

essential. Work can be done in the future to look at ways to further refine strategies for coding qualitative studies and, as noted by Dixon-woods et al. (2000), “greater use of methods of consensus decision and inter-rater reliability, could extend and improve confidence with which subjective judgements are made” (p. 131). More than a decade ago the NHS Centre for Review and Dissemination (1996) pointed to the lack of published examples of systematic reviews using good qualitative analysis, that challenge continues.

A general lack of methodological rigor was also observed, with generally low levels of internal validity, particularly in the case of Canadian based reports where just under 60% of reports were identified as having low levels of validity and another 29% being classified as moderate to low. Although major electronic databases of published peer-reviewed research literature as well as individual contact with researchers working in the field were accessed during initial phases of this review to identify relevant studies, Canadian based evaluations frequently emerged in the grey literature.

Whether these reports and studies were published as part of local school/board level initiatives, lacked sufficient budgeting for summative evaluation and results dissemination, were undertaken primarily as internal information gathering activities, or not, the overall methodological quality of studies was unacceptable. Documents such as these are used routinely to allocate funding, establish public credibility and determine future instructional and curricular initiatives, basing decisions on research, which, in far too many cases was poorly designed or at the very least poorly reported, is not conducive to the kinds of “accountability” structures and “evidence-based practice” philosophy being espoused across the country.

How Have Programs Been Designed and Delivered?

Beyond issues of statistical significance, this review also advances the idea that we should develop and implement programs that not only demonstrate efficacy in the experimental context but also effectiveness and generalizability in other situations. Applying randomization tends to strengthen internal validity, but if the experimental situation cannot be replicated in the real world, can we still rely on those results? Clark (1989) advocated for looking not only at how instruction is *designed* (developing a blueprint of generalizations about effective principles) but also how it is *developed* (addressing ‘local’ issues pertaining to implementation) and this review highlights the complimentary nature of both kinds of information.

Descriptions of program participants and instructional interventions (e.g., content of academic coursework, topics covered during life skills training, specific objectives or competencies targeted through career preparation work) were severely lacking and in most cases precluded any attempts for study replication based on reported descriptions. This issue is of particular concern in Canada because we see distinct and statistically significant differences in dropout rates across urban/rural demographics, among male and female students, among various provinces in general and across ethnic groups specifically (De Broucker, 2005). Information relating to where programs under review here fit across those dimensions would have gone a long way to illuminating issues for replicating “successful” programs in other contexts and understanding how local issues such as funding concerns, access to specialists and practitioners, cultural and ethnic practices, and economic realities play out in programs across the country.

In terms of the content of interventions reviewed here, most programs included multiple instructional mechanisms with 74% of programs incorporating at least two program components (e.g., health, life skills, career prep, academic support, cultural/spiritual enrichment, parental development, or vocational training) and 32% of programs incorporating 4 or more components. This review did not find substantial differences in terms of program components or instructional strategies (e.g., culturally responsive teaching, self-expression, group work, service learning, etc.) among the highest and lowest performing groups. This is consistent with other reviews (e.g., What Works Clearinghouse, 2008; Lehr, 2003) which have found it difficult to disentangle specific strategies or approaches found in all high performing and low performing programs. Lack of detail in reporting program implementation makes it difficult to determine if this is an issue not so much of whether a particular strategy or component was present but how it was translated in practice. Programs reporting negative findings could mean there were fundamental flaws in program design but they might also reflect deficiencies in how programs were implemented. No matter how well designed or constructed a program is, if it is implemented poorly, lasting and strong effects may prove elusive. For example, for 42% of the programs there was insufficient or missing data on general structure of the intervention schedule (i.e., did students have individualized timetables, was there a regularly scheduled class/program time, was the program set up as a drop-in centre? could students access support during out-of-school hours?, etc). In addition initial attempts at coding the extent to which one component was focused on compared to all others was not discernable within primary studies. A recent report by What Works Clearinghouse in which they present practical recommendations

for educators based on empirical research evidence made the following observation about dropout prevention programs:

“Dropout prevention interventions almost always include multiple components. This bundling of components presents challenges when reviewing levels of evidence... because evidence of the impact of specific intervention components on dropping out cannot formally be attributed to one component of an intervention” (p. 1)

Descriptions of the degree to which programs were implemented with integrity to the original design and the procedures that were in place to determine fidelity to those designs were absent in the studies reviewed here. Understanding not just IF programs work globally but who is and is not benefiting from these interventions specifically would help to establish differentiated instructional and program designs that meet specific needs of the learner. Rothwell & Kazanas (2004) argued that one of the micro trends influencing the design process at the moment is focusing on the “performance setting not the instructional setting” (p, xxxii) and re-examining what can be done to surround performers, or, in this case, learners with what they need to perform at the time they need it through a variety of support systems. This is perhaps especially relevant in the case of dropout interventions. The focus of interventions in this review tended to be on effecting change in the student not contextual features such as peers, family, school climate, etc., which is disconcerting in some respects when the prevailing theoretical perspective is for a system oriented approach to working with students at-risk (Christenson & Anderson, 2002). It is reasonable to expect that more interventions will begin to focus on non-instructional influences over learning and if evaluations of those programs are linked explicitly to enrolment and graduation rates we may be able to better determine the effects of those factors outside of the learner that may influence student performance.

Understanding how instructional programming relates to the broader support services is critical for determining “what works”. Although programs could be broadly categorized as involving “tutoring” or “parental involvement”, the implementation and design of those individual components varied widely across programs. Take, for example, one program studied by Hayward (1995) and one by Volpe (1998); each incorporated a vocational component of instruction. In the program reported by Volpe the description of that component was “ by using a partial timetable, student are required to attend classes within the school for part of the day, while the other portion is devoted to on-site training with a mentor in the community” (p. 27). In the program described by Hayward the vocational component looked quite different where students took courses at the “vocational-technical centre, entirely through CAI....were eligible to participate in the Industrial/Community-based Education program, which permitted students to work in the community once they had completed 50% of their vocational competencies and, while working, to attend school only one day a week” (p. 4-27). While both incorporate a vocational element the extent to which those instructional decisions relate to broader program goals and other elements of the instructional system are very different.

In the future researchers need to provide more description of the instructional and personal characteristics of these programs, the sequencing of instruction, the structure of activities and assessment (e.g., what kinds of tests were used to determine final grades?) the role and extent of support staff involvement (e.g., what roles and responsibilities did “student support workers” have?), the content of academic coursework (e.g., which textbook, CAI system, instructional materials were used?), the individual differences among students, etc. Determining the features of programs/participants that may explain

success can only happen if we are able to explore what and how program elements were implemented.

This review sought to increase awareness about existing research, particularly in Canada, on dropout prevention programs and to provide decision-makers with information on existing evidence of effectiveness. What has been learned through this review is that Canadian research to date has generally lacked consistency in measuring effectiveness and in applying standards of methodological rigor to research and evaluation endeavours. While there is a general consensus among programs that targeted interventions are having a positive impact on dropout rates, both statistical and descriptive evidence suggest caution in interpreting the impact these programs are having and that additional analysis and evaluation of programs, especially those that report on dropout and graduation rates, are needed to determine the extent to which program effects can be maintained over time and improved.

This review also sought to inform the discourse on what we define as “success” and how we might strategically plan and manage resources based on research evidence. The review highlighted the need for developing and applying consistent measures for computing “dropout rates” and for addressing the challenges inherent in case study research designs which often hamper our ability to generalize findings and conduct more powerful tests of intervention effects.

Finally, this review attempted to broadly explore factors which may moderate the effects of dropout prevention programs. Insufficient reporting of data for analyzing effects generally and implementation program features specifically did not give a clear

outline on how to develop efficacious and efficient programs. Although there have been, and continue to be many efforts undertaken to address the issue of early school leaving in Canada and abroad over the last few decades dropout rates continue to be part of the educational discourse and a challenge for instructional designers and policy makers.

Application for Practitioners

In systematically reviewing the available literature on dropout prevention programs in the Canadian context it has become quite clear that this literature has been ineffective in helping decision-makers transfer research knowledge into practice. As noted earlier the call for more methodologically rigorous research and more detailed reporting of implementation has in a general sense not been translated into action. This suggests at least three possibilities. One, that the message going out from researchers in the field about the need for standardized measures of dropout, reporting of the instructional environment and activities etc. has not been disseminated to decision-makers. Second, that the message has been heard but there has been a problem with implementing recommendations and building consensus around those recommendations. A third possibility might be, as Levin (1992) suggested

“our strategies rest on a simplistic and often inaccurate concept of dropping out, and a naïve faith in schooling as the way to success...The literature indicates clearly that, for many students, leaving school is a perfectly rationale thing to do... For example, dropping out appears to be related to opportunities for employment and prevailing wages...where more schooling is unlikely to lead to better employment (for example on aboriginal reserves with 80% or 90% unemployment), students deduce correctly that more schooling is not a good investment... overt political action is only part of the picture, and in some ways is the easiest part with which to grapple. Institutional inertia, protection of mandates, problems with provider capture, lack of imagination or of will are all part of the problem” (p. 262-267).

Perhaps part of this issue is that “research on managerial and policy decision making has taught us that research in the form of “ideas” not “data”, most influences decision-making” (Lavis, Robertson, Woodside, McLeod & Abelson, 2003). Finding the balance between “telling decision makers about solutions” (Lavis et al., p.224) and working to enlighten decision-makers about an issue so they can incorporate those findings within their own contexts is difficult. Decision makers are looking for specific advice on how to build instructional programs that decrease the number of students leaving school before completion and researchers are looking for the kinds of evidence from which to build a credible argument for which strategies are statistically more likely to produce that desired effect. Ultimately this project was in pursuit of opportunities to capture research impacts (e.g., to present to decision-makers “what works” in the Canadian dropout program context) as well as to increase awareness, knowledge and attitudes towards dropout prevention programming, instructional design and implementation issues. Although this review falls short of being able to provide an instructional “ideal” or directives on which options should be selected in developing these initiatives the project tells us a great deal about what we can be doing collectively to improve the quality of the evidence we use to inform our practice.

Rethinking the Existing System

When we look at dropout rates across the country (depending on the formulas used to aggregate that data) 1 in 10 Canadians 20-24 years of age were neither attending school, nor had a high school diploma. This statistic reported by STATCAN does not however, account for students who dropped out before finishing grade 12 but were able to complete a GED or high school equivalency program, so it is reasonable to assume that the number

of students dropping out before grade 12 would be higher. Males reported higher levels of dropping out from females, some estimates put the rate of graduation among aboriginal students at less than 50% and rural communities on average face nearly double the rate of dropout from their urban counterparts (17% rural, 9% urban). One might ask then, given such significant numbers, whether the existing system is in fact “working well for most students”, with only a few “at-risk” learners requiring “support services” to graduate. If even the most conservative estimate is that one in 10 do not graduate, and assuming that at least some at-risk student do not drop out so the number of at-risk student would be higher, does it even make sense to talk about an at-risk group that is nationally 10-20 % or more of the total student body? Perhaps thinking about the issue of dropouts as an issue of activity within the entire schooling system makes more sense, one that is important and impacts on all students (Levin, 202). Findings in this review, as with other recent reviews (e.g., Shannon & Bylsma, 2005), support the idea that a fundamental level what most dropout interventions do is (a) provide more enriched and intensive academic support; (b) offer a more personalized instructional process; and (c) provide content and experience more relevant to post-school situations students will be facing.

Instead of focusing solely on creating safety nets and targeted interventions for student at-risk of dropping out more could be done to think about how instruction within the mainstream classroom and the instructional setting provided within the school might provide performance support to all learners. Rummler & Brahm (1988), in discussing the systems view of human performance argued that we need an alternate view to the vacuum view of performance, that we need to think about how individuals operate and learn within the context of performance *systems* and that we can work to manage the

variables within the system to improve performance. For example, schools and boards need to examine how regulations and policies relating to absenteeism and behaviour are implemented (Shannon & Bylsma, 2005). Work can be done to better track what the causes of suspensions and expulsions are, whether policies are fairly and equitably applied, what alternatives exist for dealing with disciplinary and attendance issues there could be that keep students at school instead of removing them and thus having them have an even greater challenge for staying on-track.

Questions for stakeholders:

1. How are concepts like life-skills, career prep and vocational education integrated into the curriculum to the benefit of all students?
2. How are instructional activities implemented to allow for flexibility, student choice and self-direction?
3. What kind of training, professional development and support do teachers need to implement best-practices relating to classroom instruction?
4. Where in a students' school career do they find access to advice and guidance about academic and employment opportunities?
5. Do school structures and policies take into consideration part-time employment demands and needs of students?
6. What opportunities for remedial and in-class support are available to students to "get back on track" by recapturing lost credits through other means, receiving credit for work completed or for developing proficiency in key academic areas?

7. How do policies promote and accommodate students' re-entry to school after dropping out?
8. What happens to students who are absent for extended periods of time due to financial or personal crisis, incarceration, illness or pregnancy?

Routinely Examining and Reporting on Relevant Data

Schools and school boards need to systematically evaluate *and* report on programs and practices both to their own local communities and stakeholders but also within the broader educational community in Canada. In a recent systematic review by What Works Clearinghouse (2005) a panel of reviewers found that none of the studies identified for inclusion in their review directly evaluated the effect of using data on staying in school, but that nonetheless that “the effectiveness of targeted and school wide interventions... will depend on the extent to which they re based on an accurate assessment of the dropout problem” (p.12). The volume of material discussing strategies for dropout prevention is staggeringly voluminous and yet the discussion of the impacts (i.e., reduction of dropout rates) and even sometimes of effectiveness (i.e., measure of attitude or behaviour) are not only underreported but often disregarded. This leads to faster turnover of programs, instable funding for existing programs and the possibility that “pilot” projects are being duplicated because past implementations were not properly evaluated. Further, discussion of program costs, sustainability, durability of effects over time, required resources and implementation challenges all need to be more methodically and comprehensively reported. Collecting data without analyzing it and applying it does little to change existing practices that mediate students' chances of graduating

Moreover, regularly analyzing student data is a critical step for determining the population of students you are most interested in supporting, both the scope of the population (i.e., some schools may have a very large proportion of students at-risk of dropping out, other not as much) and the specific students who should be considered for extra services or supports (What Works Clearinghouse, 2008). Rothwell, Hohne & King, 2007, p. 182-183) argued the emerging trend in public organizations and government agencies is in re-examining their processes, structures, approaches and goals in an effort to improve overall outcomes and performance. A key characteristic of that process will be identifying individuals who have specific needs, striving to meet those needs, measuring performance and publically reporting on progress to help assure appropriate transparency and accountability.

Due to the applied nature of intervention research we may not always be able to apply random sampling but what has been widely advocated within the review literature and supported by the results of this review is the need to improve the number and the methodological quality of studies being published and undertaken (e.g., Lehr et al., 2003). As pointed out earlier that challenge may require even greater effort within the Canadian context due to lack of federal education oversight and no current national organization providing direction on research in this area. Although Canadian and American education systems have many similarities, we are not experiencing the same rates of students dropping out or the historical stagnation of rates, resistant to intervention, being observed in American school systems. Further, we do not have the same access to resources or funding opportunities as our counterparts do which may make it more difficult to engage in large-scale assessments of interventions. Improving

existing procedures for collecting and reporting on data from within and across our own communities would help us better understand intervention models that are working here and why.

One place for practitioners and stakeholders to start would be to explore existing guidelines for program evaluation. In addition to standards and reporting guidelines suggested by What Works Clearinghouse and the National Dropout Prevention Centre, several researchers have compiled useful handbooks and guidebooks for conducting evaluations on dropout prevention programs. Jay Smink (1992) has done extensive work with the National Dropout Prevention Center in developing tools for program staff to more effectively design, develop, implement and improve their dropout prevention programs. The “Evaluation Handbook: Guidelines For Evaluating Dropout Prevention Programs” which he co-wrote not only walks program staff through the steps of conducting rigorous evaluations but also provides sample data collection instruments and directions on writing up the evaluation report so that other researchers can glean as much information as possible about why programs were successful or not. Other researchers (e.g., Moberg, 1984; Little, DuPree, & Deich, 2002) have also published reference resources for practitioners to help streamline evaluation procedures and make evaluation more practical, usable and systematic. Finally, research centres like the National Dropout Prevention Center/Network (NDPC/N) provide models, case examples and standard reporting formats on their website that can guide evaluators in decision-making during the evaluation and reporting of program outcomes.

Questions for stakeholders:

1. Are graduation, enrolment, and attendance rates defined and measured consistently within classrooms, across schools, boards, provinces?
2. Does reporting on program interventions include analysis of implementation and process variables (i.e., how features within the context might be acting as barriers or enhancers to program effects)?
3. Aside from psychological, social, behavioural, familial or attendance outcomes, are all initiatives targeting students at risk of dropping collecting and analyzing data on graduation rates or enrolment status?
4. Who is overseeing the evaluation of dropout prevention interventions, both formatively and summatively? Is there potential bias or conflict of interest in that position?
5. How are results being reported to students, parents, teachers, community members, research agencies and educational agencies and organizations outside of your jurisdiction?
6. Are students being appropriately identified early in their school careers in order to provide them intervention support as early as possible?
7. Is longitudinal data being collected and analyzed that follows students through school to graduation or dropout and even beyond? Are there mechanisms to account for school or board transfers?

8. Is sufficient funding provided or research partnerships established to apply more robust research methodologies to program evaluations to increasing order to increase the generalizability of findings on program effectiveness?

Conclusion

Can we take the programs reviewed in this study that demonstrated some level of success and replicate them as models of best-practice? Probably not. Does this review provide substantiated evidence that specific instructional and program design features resulted in more significant gains in graduation rates? No. The pattern of evidence did however point to some ways that stakeholders could work towards improving effective evaluation practices in their own contexts and re-examine issues of risk both as a function of the learner and the larger instructional context; including school policies, practices, the role of the community, or current economic and labour market climate.

It is sobering that even though the “dropout crisis” has been part of the Canadian educational discourse since the 1960’s there are still significant portions of our national and regional populations that face real challenges in completing basic education. Despite unparalleled opportunities provided through technology and the advancement in research methodology for collecting, managing, analyzing synthesizing and sharing information on student performance within these interventions and programs, the body of evidence to date is marginally informative about where we have been and where we could go from here. If evaluation of programs continues to be driven by immediate concerns and for the sole purpose of providing internal feedback about funding decisions, and does not seek to more broadly engage with and communicate with the research community and other

practitioners outside the local context, then we will continue to lack the volume of research needed to provide more direct instruction on best-practices.

Perhaps here in lies the opportunity for the field of human performance technology to provide a framework for transforming the way we think about and in turn address student performance particularly when it comes to interventions designed to improve program outcomes for those at risk of dropping out. The plans put forward by human performance technologists like Rothwell (1996) sound brilliantly applicable in print:

“ a systematic process of discovering and analyzing important human performance gaps, planning for future improvements in human performance, designing and developing cost-effective and ethically justifiable interventions to close performance gaps, implementing the interventions, and evaluating the financial and nonfinancial results”

That said, many of the anxieties currently being voiced about “evidence based practice” and the growing focus on “accountability” are based on very strong values about what school means, whose interests it is for, who will design and engage in it and who will benefit from it. While this review has provided some unique examinations of the instructional and programmatic design features of dropout interventions in Canada over the last decade and a half, the general findings reaffirm the prevailing view of the last 20 years- that a more systematic analysis, implementation, evaluation and publication of results on interventions to reduce dropout rates is required. It may be that at this point in the evolution of intervention research on dropout prevention that the issue is not so much building a consensus on what should be done, but on deciding whether there is the political will and creative insight to bridge the gap between rhetoric and practice.

References

- Adams, M. (2003). *Fire and ice: The United States, Canada and the myth of converging values*. Toronto: Penguin Canada.
- Anisef, P., & Andres, L. (1996). Dropping out in Canada. The construction of a crisis? In D. Kelly & J. Gaskell (Eds.), *Debating dropouts* (pp. 84-100). New York: Teachers College Press
- Aron, L., & Zweig, J. (2003). Educational alternatives for vulnerable youth. Washington, DC: The Urban Institute. Accessed March, 2006 from:
http://www.urban.org/UploadedPDF/410898_vulnerable_youth.pdf
- Bair, C. R. (1999, November). Meta-synthesis. Paper presented at the Annual Meeting of the Association for the Study of Higher Education, San Antonio, Texas. (ERIC Document Reproduction Service No. 437866)
- Barron, Jennie. Inner-city outdoor ed. At Wilma's place. (1995). *Pathways: The Ontario Journal of Outdoor Education* ., Nov-Dec ; 7(7):14-17; ISSN: 0840-8114.
- Barton, P. (2005). One-Third of a Nation: Rising Dropout Rates and Declining Opportunities. Educational Testing Service Policy Information Report. Accessed November 2005:
<http://www.ets.org/research/pic/onethird.pdf>
- Bernard, R.M., Abrami, P.C., Lou, Y., Borokhovski, E., Wade, A., Wozney, L.,Wallet, P.A., Fiset, M. & Huang, B. (2004). How does distance education compare to classroom instruction? A meta-analysis of the empirical literature. *Review of Educational Research*, pp. 379-439
- Bickel, W. E. (1991). Images of effective dropout programs. *Contemporary Psychology*, 36, 73-73.
- Bloom, M. (1991). Profiles of partnerships: Business-education partnerships that enhance student retention. . Ottawa: Conference Board of Canada
- Boaz, A., Ashby, D., & Young, K. (2002). Systematic reviews: what have they got to offer evidence based policy and practice? ESRC UK Centre for Evidence Based Policy and Practice, Queen Mary, University of London. Retrieved July 1, 2002 from
<http://www.evidencenetwork.org/cgi-win/enet.exe/biblioview?538>
- Bowlby, G. & McMullen, K. (2005). Provincial dropout rates – trends and consequences. *Education Matters: Insights on Education, Learning and Training in Canada*, 2(4). Statistics Canada Catalogue number 81-004-XIE. Accessed March, 2006 from:
<http://www.statcan.ca/english/freepub/81-004-XIE/2005004/drop.htm>
- Brown, R. S. (1994). A two-year evaluation of the change your future program at the Toronto board of education. 1991-1993. no. 208. Canada; Ontario: Toronto Board of Education Research Dept.
- Bushman, B. J. (1994). Vote counting procedures in meta-analysis. IN H. Cooper & L. Hedges (Eds.) *The handbook of research synthesis* (pp.193-214). New York: Russell Sage.
- Canadian Council on Learning [CCL] (2005). Good news: Canada's high school dropout rates are falling. Lessons in Learning Report. Accessed March, 2006 from: <http://www.ccl-cca.ca/english/resources/carnet.asp>

- Christensen, S. L., Sinclair, M. F., Lehr, C. A., & Hurley, C. M. (2000). Promoting successful school completion. In K. Minke & G. Bear (Eds.), *Preventing school problems—promoting school success: Strategies and programs that work*. Bethesda, MD: National Association of School Psychologists.
- Clark, R. E. (1989). Current progress and future directions for research in instructional technology. *Educational Technology Research & Development*, 37(1), 57-66.
- Cobb, B., Sample, P., Alwell, M., & Johns, N. (2005). Effective interventions in dropout prevention: A research synthesis. National Dropout Prevention Center for Students with Disabilities. Clemson, SC. Accessed March, 2006 from: <http://www2.edc.org/ndpc-sd/cognitive/report.pdf>
- Collinson, V., Fedoruk, T., & Conley, S. (2001, September). Organizational learning: Theory and applications for schools and school systems. Presented at the 26 ATEE conference in Stockholm. Retrieved July 15th at URL: http://www.lhs.se/atee/papers/Concurrent_3_1.doc
- Cook, D. J., Mulrow, C. D., & Haynes, R. B. (1997). Systematic reviews: Synthesis of best evidence for clinical decisions. *Annals of Internal Medicine*, 126, 376-380.
- Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design & analysis issues for field settings*. Boston: Houghton Mifflin Company
- Cook, T. D., & Leviton, L. C. (1980). Reviewing the literature: A comparison of traditional methods with meta-analysis. *Journal of Personality*, 48, 449-472.
- Cook, T. D., Cooper, H., Cordray, D.S., Hartmann, H., Hegdes, L.V., Light, R.J., Louis, T.A. & Mosteller, F. (1992). *Meta-analysis for explanation*. NY: Sage
- Coombs, R. & Falletta, J. (2000). The targeted evaluation process; A performance consultant's guide to asking the right questions and getting the results you trust. Alexandria, VA: American Society for Training and Development.
- Coutts, R.; Chinien, C. A.; Boutin, F., Letteri, C., Cap, O., & Porozny, G. (1995) Dropout prevention through a cognitive based instructional system: Why are we not all born smart? Final report; 172.
- Cooper, H. & Hedges, L.V (1994). *The handbook of research synthesis*. New York: Russell Sage Foundation
- Cooper, H. M. (1984). *The integrative research review*. Beverly Hills, CA: Sage
- Coscarelli, B. (1988). Performance improvement quarterly and human performance technology. *Performance Improvement Quarterly*, 1(1), 2-5.
- Currie, S.; Foley, K., Schwartz, S., & Taylor Lewis, M. (2001). Bladerunners and picasso cafe: A case study evaluation of two work-based training programs for disadvantaged youth.
- De Broucker, P. (2005). Without a Paddle: What to do about Canada's young dropouts? Canadian Policy Research Networks. Ottawa. Accessed March, 2006 from : <http://www.cprn.org/en/doc.cfm?doc=1332>
- Dixon-Woods, M. , Bonas, S., Booth, A., Jones, D.R., Miller, T., Shaw, R.L., Smith, J., Sutton, A., & Young, B. (2006). How can systematic reviews incorporate qualitative research? A critical perspective. *Qualitative Research* 6: 27-44

- Dixon-Woods, M., Bonas, S., Booth, A., Jones D.R., Miller, T., & Shaw, R. L., et al. (2006). How can systematic reviews incorporate qualitative research? A critical perspective. *Qualitative Research*, 6(27), 27-44.
- Dixon-Woods, M., Fitzpatrick, R., & Roberts, K. (2001). Including qualitative research in systematic reviews: problems and opportunities. *Journal of Evaluation in Clinical Practice* 7: 125-133.
- Dynarski, M. (2000) How can we help? What we have learned from federal dropout-prevention Programs. Paper presented at the School Completion in Standards-Based Reform: Facts and Strategies Workshop, National Research Council, Washington, DC (July 17). Available: Mathematica Policy Research
- Dynarski, M. (2004). Interpreting the evidence from recent federal evaluations of dropout-prevention programs: The State of scientific research. In Orfield, G. (Ed.) *Dropouts in America: Confronting the Graduation Rate Crisis*. Cambridge, MA: Harvard Education Press (p. 255-268).
- Dynarski, M., & Gleason, P. (2002). How can we help? What we have learned from recent federal dropout prevention programs. *Journal of Education for Students Placed At-Risk*, 7(1), 43-69.
- Eckhardt, K.W., & Ermann, M.D. (1977). Social research methods. New York: Random House
- Epstein, J.L., & Sheldon, S.B. (2002). Present and accounted for: Improving student attendance through family and community involvement. *Journal of Educational Research*, 95(5), 308-318.
- Estabrooks, C., Field, P A., & Morse, J. (1994). Aggregating qualitative findings: An approach to theory development. *Qualitative Health Research*, 4(4), 503-511.
- Eyesnck, H. J. (1984). Meta-analysis: An abuse of research integration. *The Journal of Special Education*, 18(1), 41-59.
- Fitzgerald, J. (1995). English as a second language learners' cognitive reading processes: a review of research in the United States. *Review of Educational Research*, 65(2), 145-190.
- Ford, R., Gurr, S.; Ivry, R.J., & Taylor-Lewis, M. (1999). (Social Research and Demonstration Corporation). Transitions: Programs to encourage British Columbia students to stay in school, June; SRDC Working Paper Series 99-01.
- Galbraith, J. (Winter, 1982). Designing the innovating organization. *Organizational Dynamics*, 5-24.
- Glass, G. V. (1976). Primary, secondary and meta-analysis of research. *Educational Research*, 5(10), 3-8.
- Glass, G. V., McGaw, B., & Smith, M. L. (1981). *Meta-analysis in social research*. Beverly Hills, CA: Sage.
- Gleason, P. & Dynarski, M. (1998). *How can we help? What have we learned from evaluations of federal dropout prevention programs*. Princeton, NJ: Mathematica Policy Research, Inc
- Goldschmidt, P., & Wang, J. (1999). When can schools affect dropout behavior? A longitudinal multilevel analysis. *American Educational Research Journal*, 36, 715-738.

- Gordon, Joan M. (2004). Looking at high school dropout problems from students' perspectives: Finding a solution: McGill University (Canada); Ph.D.
- Grannis, J. (1991). Dropout prevention in New York City: A Second chance. *Phi Delta Kappan*, 73(2), 143-149.
- Harless, J. (1998). *The Eden conspiracy: Educating for accomplished citizenship*. Wheaton, IL: Guild V Publications
- Hedges, L. V. & Olkin, I. (1985). *Statistical methods for meta-analysis*. Orlando, FL: Academic Press.
- Hedges, L. V., & Pigott, T. D. (2001). The power of statistical tests in meta-analysis. *Psychological Methods*, 6, 203-217.
- Hedges, L. V., Shymansky, J. A., & Woodworth, G. (1989). A practical guide to modern methods of meta-analysis. [ERIC Document Reproduction Service No. ED 309 952].
- Hopewell, S., Clarke, M., & Mallet, S. (2005). Grey literature and systematic reviews. In H. Rothstein, A. Sutton & M. Borenstein (Eds.), *Publication bias in meta-analysis prevention, assessment and adjustments* (pp. 49-72). West Sussex, England: Wiley and Sons Inc
- Hossler, D., & Scalese-Love, P. (1989). Grounded meta-analysis: A guide for research synthesis. *Review of Higher Education*, 13(1) 1-28.
- Human Resources and Social Development Canada (2000). Dropping out of high school: definitions and costs. Retrieved: <http://www.hrsdc.gc.ca/en/cs/sp/hrsd/prc/publications/research/2000-000063/page05.shtml>
- Hunter, J. & Schmidt, F. (2004). *Methods of meta-analysis: correcting error and bias in research findings* (2nd. Ed.). Sage: Thousand oaks
- Jensen, L., & Allen, M. (1994). A synthesis of qualitative research in wellness-illness. *Qualitative Health Research*, 4(4), 349-369.
- Jordan, A. (2006). Background research paper for a national dialogue on students at risk. Report Prepared for the National Dialogue on Students at Risk, Vancouver, BC February 26-28, 2006
- Jordan, W. J., J. M. McPartland & J. Lara (1999). Rethinking the causes of high school dropout. *The Prevention Researcher*, 6 (3), 1-4.
- King, A. (2005). Double cohort study: Phase 4 report (for the Ontario Ministry of Education) Retrieved March 2006 from: <http://www.edu.gov.on.ca/eng/document/reports/phase4/index.html>
- Kratochwill, T. R., Stoiber, K. C., Christenson, S. L., Durlak, J. A., Levin, J. R., Waas, G., et al. (2002). Procedural and coding manual for identification of evidence-based interventions. Task Force on the Evidence Based Interventions in School Psychology, Sponsored by Division 16 of the American Psychological Association and Society for the Study of School Psychology. Accessed March, 2006 from: http://www.sp-ebi.org/documents/_workingfiles/EBImanual1.pdf
- Laiken, M. (2002). Models of organizational learning: Paradoxes and best practices in the post industrial workplace. ERIC Clearinghouse on Adult, Career and Vocational Education; NALL Working Paper #25-2001

- LaPalombar, J. (2001). Power and politics in organizations: Public and private sector comparisons. In M. Dierkes, A. Berthoin Antal, J. Child, & I. Nonaka (Eds.), *Handbook of Organizational Learning and Knowledge* (pp. 557-581). Oxford: Oxford University Press.
- Lawler, Edward, III. (1996). *From the ground up: Six principles for the new logic corporation*. San Francisco: Jossey-Bass.
- LeCompte, M. D., & Dworkin, A. G. (1991). *Giving up on school: Student dropouts and teacher burnouts*. Newbury Park, CA: Corwin Press Inc.
- LeCompte, M.D., & Gooetz, J. P. (1982). Problems of reliability and validity in ethnographic research. *Review of Educational Research*, 52(1), 31-60.
- Lehr, C. A., Hansen, A., Sinclair, M. F., & Christenson, S. L. (2003). Moving beyond dropout prevention to school completion: An integrative review of data-based interventions. *School Psychology Review* 32(3), 342-364.
- Lehr, C. A., Johnson, D. R., Bremer, C. D., Cosio, A., & Thompson, M. (2004). Essential tools: Increasing rates of school completion: Moving from policy and research to practice. Minneapolis, MN: University of Minnesota, Institute on Community Integration, National Center on Secondary Education and Transition. Accessed March, 2006 from: <http://www.ncset.org/publications/essentialtools/dropout/dropout.pdf>
- Levin, B. (1992). Dealing with dropouts in Canadian education. *Curriculum Inquiry*, 22, 257-270
- Levin, M., & Sanger, M. B. (1994). *Making Government Work: How Entrepreneurial Executives Turn Bright Ideas Into Real Results*. San Francisco: Jossey-Bass.
- Light, R. J., & Pillemer, D. B. (1984). *Summing up: The science of reviewing research*. Cambridge, MA: Harvard University Press.
- Lincoln, Y. S., Guba, E.G. (1985). Processing naturalistically obtained data. In Y. S. Lincoln and E. G. Guba, *Naturalistic Inquiry* (pp. 332-356). Beverly Hills, CA: Sage
- Lipsey, M. W., & Wilson, D. B. (2001). Practical meta-analysis. Applied social research methods series (Vol. 49). Thousand Oaks, CA: SAGE Publications.
- Little, P., DuPree, S., & Deich, S. (2002). Documenting progress and demonstrating results: Evaluating local out-of-school time programs. Cambridge, MA: Harvard Family Research Project and The Finance Project. Available online at www.gse.harvard.edu/hfrp/content/projects/afterschool/resources/issuebrief3.pdf.
- Long, A. F., Godfrey, M., Randall, T., Brett, A., & Grant, M. J. (1999) Developing evidence based social care policy and practice: final report. Part 3: feasibility of undertaking systematic reviews in social care University of Leeds, Nuffield Institute for Health, 71-75 Clarendon Road, Leeds LS2 9PL, and University of Salford, Health Care Practice R&D Unit
- Lucas, W. A. (1974). *The case survey method: Aggregating case experience*. Santa Monica, CA: Rand.
- MacLean, D.A., & Janzen, H. L. (1994). A framework for keeping students in school. *Canadian Journal of School Psychology*, 10(1), 54-61.

- Mager, R. F. & Pipe, P. (1997). *Analyzing performance problems*. Atlanta: The Center for Effective Performance
- Marrelli, A. F. (2005). The performance technologist's toolbox: Literature reviews. *Performance Improvement*, 44, 40-44.
- Matt, G. & Cook, T. (1994). Threats to the validity of research synthesis. In Cooper, H. and Hedges, L (eds.) *The Handbook of research synthesis*. New York: Russell Sage Foundation
- McGaw, B. (1997). Meta-analysis. In J. P. Keeves (Ed.), *Educational research, methodology, and measurement: an international handbook* (2nd ed.), pp. 371-380. Cambridge: Elsevier Science Ltd.
- Merrill, M. D., Reigeluth, C. M., & Faust, G. W. (1979). The instructional quality profile: A curriculum evaluation and design tool. In Procedures for Instructional Systems Development. In H. A. O'Neil, Jr. (ed.) *Procedures for Instructional Systems Development*. New York: Academic Press.
- Milak, D (2001). Windows of opportunity for the at-risk children preschool early intervention. Master's Project Brandon University.
- Miles, M., & Huberman, M. (1994). *Qualitative data analysis* (2nd ed.). Thousand Oaks, CA: Sage.
- Miller, N., & Pollock, V.E. (1995). Use of meta-analysis for testing theory. *Evaluation and the Health Professions*, 18, 370-392.
- Moberg, D.P. (1984). Evaluation of prevention programs: A basic guide for practitioners. Madison, WI: Wisconsin Clearinghouse.
- Morgan, G. A., Gliner, J. A., & Harmon, R. J. (2006). Understanding and evaluating research in applied and clinical settings. Mahwah, N J: Lawrence Erlbaum Associates.
- Morris, S., Pawlovich, W., & McCall, D. (1991). Evaluating the effectiveness of school dropout prevention strategies: some suggestions for future research: final report. Toronto, Ontario: Canadian Education Association
- Natriello, G. (Ed.). (1986). School dropouts: Patterns and policies. New York: Teachers College Press. (ED 291 811)
- Neil, D. & Wykowski, T. (1999). *Beyond Reductionism: Gateways for Learning and Change*. Boca Raton: St. Lucie Press.
- New Approaches, New Solutions - Fostering Success for Secondary School Students in Disadvantaged Areas (2002) QUEBEC. Retrieved March 2006 , from: <http://www.qesnrecit.qc.ca/reform/success/docs/NANS1.pdf>
- Noblit, G. W, & Hare, R D. (1988). *Meta-ethnography: Synthesizing qualitative studies*. Newbury Park, CA: Sage.
- Oakley, A. (2003). Research evidence, knowledge management and educational practice: Early lessons from a systematic approach. *London Review of Education* 1 (1), 21-33.
- Ogles, Benjamin M., Lambert, Michael, J., & Masters, Kevin S. (1996). Assessing outcome in clinical practice. Boston: Allyn and Bacon.

- Oliver, S., Harden, A., Rees, R., Shepherd, J., Brunton, G., Garcia, J., & Oakley, A. (2005). An emerging framework for including different types of evidence in systematic reviews for public policy. *Evaluation*, 11, 428-446.
- Pascarella, E. & Terenzini, P. (1991). *How college affects students: Findings and insights from twenty years of research*. San Francisco: Jossey-Bass.
- Paterson, B. L., Thorne, S. E., Canam, C., & Jillings, C. (2001). *Meta-Study of qualitative health research*. Thousand Oakes, CA: Sage.
- Pawson, R. (2002). Evidence-based policy: The promise of "realist synthesis". *Evaluation*, 8(3):340-358.
- Pershing, J. (2006). Human performance technology fundamentals. In J. Pershing (ed.) *Handbook of Human Performance Technology: Principles, Practices, and Potential*, 3rd Edition. San Francisco: Pfeiffer.
- Petitti D B. (1994). *Meta-analysis, decision analysis, and cost-effectiveness analysis*. New York: Oxford University Press
- Petticrew, M., & Roberts, H. (2006) *Systematic reviews in the social sciences*. Oxford: Blackwell Publishing
- Phaneuf, A. (1993). Un projet d'école alternative. *Inform-Action* 22(2), 28-30.
- Pinkus, L. (2006). Who's Counted, who's counting?: Understanding high school graduation rates. Report published by The Alliance for Excellent Education. Retrieved from WWW at http://www.all4ed.org/publication_material/reports/whoscounted_counting
- Poister, T. H. (1978). *Public program analysis: Applied research methods*. Baltimore: University Park Press.
- Price Waterhouse (1990). Qualitative Research Related to the School Leavers Questionnaire" Final Report, October 25, 1990. Ottawa: Minister of Supply and Services.
- Ragin, C. C. (1987). *The comparative method: moving beyond qualitative and quantitative strategies*. Berkeley: University of California Press.
- Reigluth, C. (Ed.) (1999). *Instructional-Design Theories and Models: A New Paradigm of Instructional Technology*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Reyhner, J. (1992). American Indians out of school: A review of school-based causes and solutions. *Journal of American Indian Education*, 31(3), 37-56.
- Robertson, H. J. (2004). Friendly fire. *Phi Delta Kappan*, 86 250-251
- Roderick, M. (1993). *The path to dropping out: Evidence for intervention*. Westport, CT: Auburn House.
- Rosenthal, B. S. (1998). Nonschool correlates of dropout: An integrative review of the literature. *Children & Youth Services Review*, 20(5), 413-433
- Rothwell W. J. & Kazanas. H.C. (2004). *Improving On-the-Job Training: How to Establish and Operate a Comprehensive OJT Program*. San Francisco, Calif.: Pfeiffer

- Rothwell, W., Lindholm, J., & Wallick, W. (2003). *What CEOs expect from corporate training: Building workplace learning and performance initiatives that advance organizational goals*. New York: Amacom.
- Rumberger, R. W. (1987). High school dropouts: A review of issues and evidence. *Review of Educational Research*, 57, 101-121.
- Rumberger, R. W. (1995). Dropping out of middle school: A multilevel analysis of students and schools. *American Educational Research Journal*, 32(3), 583-625.
- Rumberger, R. W. (2001, January). Why students drop out of school and what can be done. Paper presented at the "Dropouts in America: How Severe Is the Problem? What Do We Know About Intervention and Prevention?" conference, Cambridge, MA.
- Sandelowski, M., Docherty, S., & Emden, C. (1997). Focus on qualitative methods. Qualitative meta-synthesis: issues and techniques. *Research in Nursing and Health*, 20(4), 365-71.
- Schargel, F. P., & Smink, J. (2001). *Strategies to help solve our school dropout problem*. Larchmont, NY
- Schofield, J. W. (1990). Increasing the generalizability of qualitative research. In E. W. Eisner & A. Peshkin (Eds.), *Qualitative inquiry in education: The continuing debate*. (pp. 201-232). New York: Teachers College Press
- Scott, W. R. (1998). *Organizations: Rational, Natural and Open Systems* (4th Ed.). Upper Saddle River, NJ: Prentice Hall.
- Shannon, G.S. & Bylsma, P. (2005). Promising Programs and Practices for Dropout Prevention: Report to the Legislature. Office of Superintendent of Public Instruction. Olympia, WA. Accessed March, 2006 from: <http://www.k12.wa.us/research/pubdocs/PromisingProgramsandPractices.pdf>
- Slavin, R.E. (1986). Best-evidence synthesis: An alternative to meta-analytic and traditional reviews. *Educational Researcher*, 15 (9), 5-11.
- Smink, J. (1992). *The Evaluation Handbook: Guidelines for Evaluating Dropout Prevention Programs*. Clemson, SC: The National Dropout Prevention Center.
- Social Sciences Research Unit [SSRU] (2004). SSRU Research: 5. Evidence for Policy and Practice Information and Co-ordinating (EPPI) Centre. Accessed March, 2006 from : <http://ioewebserver.ioe.ac.uk/ioe/cms/get.asp?cid=11995>
- Somers, C. L., & Piliawsky, M. (2004). Dropout prevention among urban, African American adolescents: Program evaluation and practical implications. *Preventing School Failure*, 48(3), 17-20.
- Steinberg, A., & Almeida, C. (2004). The dropout crisis: Promising approaches in prevention and recovery. Boston, MA: Jobs for the Future, 6. Accessed March 2006, from: <http://www.edroundtable.state.in.us/meetings/2005docs/December%202005/dropoutcrisis.pdf>
- Stock, W. A. (1994). Systematic coding for research synthesis. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 125-138). New York: Russell Sage Foundation
- Stolovitch, H. D. (2000). Human Performance Technology: Research and Theory to Practice. *Performance Improvement*, 39 (4), 7-16.

- Stufflebeam, D.L. (2001). The meta-evaluation imperative. *American Journal of Evaluation*, 22 (2), 183-210.
- Sugarman, B. (1997). The learning organization: Implications for training. In Brown, S. M., & Seidner, C. J. (Eds.), *Evaluating corporate training: Models and issues*. Norwell, MA: Kluwer Academic Publications.
- Sugrue B. & Fuller J. (Eds.) Performance interventions: Selecting, implementing, and evaluating the results (pp. 219-232). Alexandria, VA: American Society for Training and Development
- Suri, H. (2000). A critique of contemporary methods of research synthesis. Retrieved July 1, 2002, from <http://www.edfac.unimelb.edu.au/insight/postscriptfiles/vol1/suri.pdf>
- Tanner, J., Krahn, H. and Hartnagel, T. (1995). *Fractured transitions from school to work: Revisiting the dropout problem*. Toronto, Ontario: Oxford University Press.
- Thorne, S. (1994). Secondary analysis in qualitative research: Issues and implications. In J. M. Morse (Ed.), *Critical issues in qualitative research methods* (pp. 95-115). Thousand Oaks, CA: Sage.
- Vitaro, F., Brendgen, M., & Tremblay, R. E. (1999). Prevention of school dropout through the reduction of disruptive behaviors and school failure in elementary school. *Journal of School Psychology*, 37, 205–226.
- Wagennar, T.C. (1987). What do we know about dropping out of high school?, *Research in the Sociology of Education and Socialization*, 7,161-190.
- West, L. (1991). *Effective strategies for dropout prevention of at-risk youth*. Gaithersburg, Md.: Aspen Publishers
- What Works Clearinghouse (2008). Dynarski, M., Clarke, L., Cobb, B., Finn, J., Rumberger, R., and Smink, J. (2008). Dropout Prevention: A Practice Guide (NCEE 2008–4025). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved from <http://ies.ed.gov/ncee/wwc>.
- Wolf, F. M. (1986). *Meta-analysis: Quantitative methods for research synthesis*. Beverly Hills, CA: Sage.
- Yeaton, W. H., & Wortman, P. M. (1984). *Issues in data synthesis*. San Francisco: CA, Jossey-Bass.

APPENDICES

APPENDIX A- SEARCH STRATEGY SUMMARY

The following is a brief description (with an example) of each database used in the search to identify relevant studies. The overall search strategy utilized terms to capture the outcome, intervention, setting and participants. When available the database thesaurus was consulted for specific terminology.

ELECTRONIC DATABASES:

ERIC (Education Resources Information Center)

Thesaurus: Academic Persistence; Dropout Prevention; Dropout Rate; Dropouts; Persistence; School Holding Power; Student Attrition; Withdrawal (Education); program evaluation, program effectiveness

Search: (dropouts or educational mortalit* or pushouts or out of school youth or at-risk or non-graduates or disaffiliates or dropout* or early school leav* or stopouts) AND prevention or program or intervention or high-school equivalency AND Canada or any of the Canadian provinces (CP:= Canada or Alberta. or Nova Scotia...etc.)

Delimiters PY>1989

Notes: Using keywords offered a more precise search, when descriptors were used there were too many irrelevant hits. Multiple combinations of the above search terms and those listed in the methods section were used to conduct the search.

PsycINFO and EBSCO Academic Search Primer

Thesaurus: school enrolment, potential dropouts, school dropouts, student attrition, student retention

Search: dropout or early-school leav* or at-risk or graduation rate or enrolment or retention AND school* AND (prevention or intervention or program) AND (Canad* or provinc*)

Delimiters: Publication type: unlimited
Language: English or French
Population: human
Publication year: 1990-2006

Notes: Multiple combinations of individual search terms were also used to carry out electronic searches.

Canadian Research Index

Thesaurus: none

Search: (dropouts or student retention or early-school leavers or student attrition AND school*)

Delimiters: Publication year: 1990-2006

Notes: Use of descriptors limited hits. Multiple combinations of individual search terms were also used to carry out electronic searches and searching the abstracts was more productive.

CBCA Education:

Thesaurus: none

Search: (dropouts or student retention or early-school leavers or student attrition AND school*)

Delimiters: Publication year: 1990-2006

Notes: Use of descriptors limited hits. Multiple combinations of individual search terms were also used to carry out electronic searches and searching the abstracts was more productive.

Digital Dissertations on ProQuest

Thesaurus: None

Search: (dropout) OR (early-school leaver) OR (student retention) AND (prevention) OR (intervention) OR (graduation rate) OR (educational mortalit*) OR (out of school youth) AND prevention or program or intervention or high-school equivalency AND Canada or any of the Canadian provinces (CP:= Canada or Alberta. or Nova Scotia...etc.)

Delimiters: Publication year: 1990-2006

Notes: Performed this search in stepwise fashion with “dropout” and synonyms first and then outcome and target population terms second.

Educational Technology Abstracts

Thesaurus: none

Search: (dropout) OR (early-school leaver) OR (student retention) OR (graduation rate) OR (educational mortalit*) OR (out of school youth) AND (prevention) OR (intervention)

Delimiters: 1995-2006

Notes: Searches were also conducted on single terms in isolation

MANUAL SEARCHES IN JOURNALS: (Scanning of Abstracts/Table of contents)

Canadian Journal of Program Evaluation (1990-2006)
Journal of Education for Students Placed at Risk (1990-2006)
The International Journal on School Disaffection (2003-2005)

WEB SEARCH TOOLS (last search December 2006)

Google, All theWeb (Fast) & Google Scholar

Google search (1): hits 2,890,000

Phrase: "dropout prevention"

All of the words: research

Any of the words: Canada, OR nova scotia, OR new brunswick, OR prince edward island, OR ontario, OR manitoba, OR saskatchewan, OR british columbia, OR north west territories, OR newfoundland, OR Alberta

Google search (2): hits 42,500

Exact Phrase: dropout prevention program

Any of the words: canada, OR nova scotia, OR new brunswick, OR prince edward island, OR ontario, OR manitoba, OR saskatchewan, OR british columbia, OR north west territories, OR newfoundland, OR Alberta

Google search (3): hits 13,300

Exact Phrase: student retention program

Any of the words: canada, OR nova scotia, OR new brunswick, OR prince edward island, OR ontario, OR manitoba, OR saskatchewan, OR british columbia, OR north west territories, OR newfoundland, OR Alberta

Notes: Since the indexing function with Google and Google Scholar are limited and there was no way to limit by date of publication this resulted in too many irrelevant hits. Multiple combinations of the search terms were used to narrow the number of references to be reviewed. Even then the number of hits was substantial so decisions were made based on the returned titles alone, not on abstracts or viewing of the actual document. Additional terms such as “report”, “annual report”, “evaluation”, “final report”, “and findings” were also used during searches to help distinguish between documents relating to a specific program and general commentary on dropout issues.

WEB SITES (search updated December 2006)

All provincial government websites:

- [Alberta Learning](#)
- [British Columbia Ministry of Education](#)
- [Manitoba Education and Training](#)
- [New Brunswick Department of Education](#)
- [Newfoundland and Labrador Department of Education](#)
- [Northwest Territories Department of Education, Culture and Employment](#)
- [Nova Scotia Department of Education](#)
- [Nunavut Department of Education](#)
- [Ontario Ministry of Education and Training](#)
- [Québec, Ministère de L'Éducation](#)
- [Saskatchewan Education](#)
- [Yukon Education](#)

In addition:

- [Society for the Advancement of Excellence in Education](#)
- [Canadian Education Association](#)
- Centre for Research on Youth At-Risk
- Canadian Evaluation Society
- LEARN QUEBEC

BIBLIOGRAPHIES/PROCEEDINGS USED FOR BRANCHING:

Bloom, M. (1991). Profiles of Partnerships: Business-Education Partnerships that Enhance Student Retention. Ottawa: Conference Board of Canada.

The Stay in School Initiative: Ideas that work. (1996) Human Resources Development Canada (FG MP 78-4/8)

O.S.S.T.F. (Ontario Secondary School Teachers' Federation) Library/Research Centre. Bibliography on Dropouts, updated February 2004
<http://www.osstf.on.ca/www/services/library/whatsnew/bibliog/DROPOUTS.pdf>

CONCEPTUAL STUDIES AND OTHER REVIEWS (EXAMPLES)

Dynarski, M. (2004). Interpreting the Evidence from Recent Federal Evaluations of Dropout-Prevention Programs: The State of Scientific Research. In Orfield, G. (Ed.) Dropouts in America: Confronting the Graduation Rate Crisis. Cambridge, MA: Harvard Education Press

Shannon, G.S. And Bylsma, P. (2005). Promising Programs And Practices For Dropout Prevention: Report To The Legislature. Office Of Superintendent Of Public Instruction. Olympia, Wa. Accessed March, 2006 From:
[Http://www.k12.wa.us/research/pubdocs/promisingprogramsandpractices.pdf](http://www.k12.wa.us/research/pubdocs/promisingprogramsandpractices.pdf)

APPENDIX B- CODING SCHEME

These codes were adapted or reused from previous reviews (e.g., Lehr, 2003). The codebook is divided into the following sections:

Section 1 – Study Identification
Section 2 - Methodological Features
Section 3 - Objectives
Section 4 - Environmental Features
Section 5 - Resources
Section 6 - Components
Section 7 - Management

Section 1 –Study Identification

- 1) Study Number (EXCEL label: “Study”): Each article (include or exclude) is given a unique three/four digit number
- 2) Finding Number (EXCEL label: “Finding”): Each finding within a study gets a two digit number starting at 01 for the first finding in each study.
- 3) Author Name (EXCEL label: “Author”): Each study is identified in the order of authorship.
- 4) Year of Publication (EXCEL label: “Yr”): Each study is identified by year of publication. In case of a dissertation and a published article. The published article will be used.
- 5) Type of Publication (EXCEL label: “Typpub”): (Describes the source of the study. A “report” is a conference paper, governmental report, university report or ED that is not a dissertation)
 1. Journal
 2. Book chapter
 3. Report
 4. Dissertation

Section 2 – Methodological Features

- 6) Treatment Duration (EXCEL label: “Durat”): (Length of time to which any one given student is exposed to the treatment. This is often equivalent to the length of the course. One semester is defined as approximately three months or three credits.)
 1. Less than one semester
 2. One semester
 3. More than one semester
 999. Missing
- 7) Research design. (EXCEL label: “Design”): (Description of research design [e.g., case study, pre-post test design, etc.]) Random assignment= random assignment of participants to two or more groups and manipulation of the independent variable; Non-random assignment = studies including a comparison group but did not randomly assign participants to treatment or control group. Pre-post= studies that include measures and comparisons of outcome data before and after the intervention using one group. Case study = in-depth study of a phenomenon from the perspective of the participants involved. Action research= self-reflective narrative undertaken by participants in order to improve their own practices. Ethnographic research= description of events that occur within the life of a group, with regard to social structures, group membership and the

interpretation of meaning. Evaluation study= process of making judgments about the merit, value or worth of educational programs (evaluation studies are often published by the government/agency funding the project or directly involved in the design of the program)

1. random assignment
 2. non random assignment
 3. pre-post
 4. case study
 5. action research
 6. ethnographic
 7. evaluation
 8. other (define)
 999. Missing
- 8) Participant Selection (EXCEL label: "select"): This is an indicator of how participants were "eligible" or "selected" to participate in the program.
1. history of academic performance
 2. attendance records
 3. teacher referral
 4. history of dropout
 5. socio-economic background
 6. behavioral or psychological issues
 7. age
 999. Missing/Other
- 9) Statistical significance. (EXCEL label: Sig): (A mathematical determination that indicates the presence of an effect that is unlikely to have resulted from chance alone. When key outcomes are statistically significant, the intervention/program is assumed to have had an effect)
- 10) Effect size. (EXCEL label: Effect) (This indicator of effectiveness measures the amount of impact attributed to the program or intervention, and is not influenced by sample size)
- 11) Durability of effects. (EXCEL label: Durab): (Evidence indicating program effects persisted over time) Studies will be coded as YES if researchers followed participants for at least 1 year after the intervention ceased.
1. Yes
 2. No
 999. Missing
- 12) Sample size. (EXCEL label: "N"): (The sample size used in the research studies is specified)
- 13) Outcome Measure (EXCEL label: "Measure"): (Describes the validity and/or reliability of the outcome measure as a function of its source. Teacher/Board initiated= regular course tests, grades assigned to students or board level instruments used for their own purposes; Research initiated= use of published/well known behavioral/cognitive etc. tests/instruments and/or use of researcher's new instrument. Independent Evaluator = instruments developed by outside agency, consultants, etc. who are not the primary researcher)
1. Teacher/Board
 2. Researcher
 3. Independent Evaluator
 999. Missing

Section 3- Objectives

- 14) Intervention Program or Strategy. (EXCEL label: “Form”): (The title of the program is listed. In cases where the intervention did not have a formal title, the type of program is listed using summary descriptors)
- 15) Background. (EXCEL label: “Back”): (History and purpose of the program is briefly described [e.g., grass-roots, versus government developed; break-away program from another initiative, etc.])
 1. government developed/run (provincial or federal or combination)
 2. non-profit community organization
 3. for-profit community organization
 4. school based program
 999. Missing
- 16) Outcome Type (EXCEL label: “Outcome”): (Describes the basic outcomes explored by the researchers as an indicator of “success”. Academic/cognitive = GPA, standardized test scores, study habits, etc. Physical presence = enrolment status, attendance, graduation rate, dropout rate, etc. Psychological = student attitudes towards learning or school, self-esteem, depression, etc.) Social Behavioral = problem behaviour, social competence, drug use, violent behaviour, etc. Support for learning = student attitude toward teachers, school climate.
 1. Academic/cognitive
 2. Physical presence
 3. Psychological
 4. Social behavioral
 5. Support for learning
 999. Other
- 17) Program goal. (EXCEL label: “Pgoal”): (e.g., prevention, intervention, school-work transition or re-entry)
 1. Early prevention
 2. Intervention for “at risk” students
 3. School-to work transition
 4. Re-entry
 999. Missing

Section – 4 Environment

- 18) Grade Level: (EXCEL label: “grade”): Population intended for the program is briefly described with regard to grade level. NOTE: if the study indicates “secondary school” code as 5; General guidelines for ages are Elementary =age 5-11, Junior High = 12-15 and high school = 16-18)
 1. Pre-school or earlier
 2. Elementary school
 3. Junior high school
 4. High school
 5. Multi-age
 999. Missing
- 19) Setting. (EXCEL label: “setting”) .The setting in which the program has been implemented is described (e.g., urban or rural) In cases where the setting is not clearly specified (i.e. rural or urban), use the measure of population size, where population greater or equal to 50,000 corresponds to urban setting; otherwise, the delivery site should be considered rural. Population figures are to be drawn from census data quoted by authoritative sources (for example, an Atlas, or an official website) or directly from a government census bureau (for example, <http://www.census.gov/Press-Release/www/2003/SF4.html>). Document the information source in the analysis.
 1. Rural

- 2. Urban
- 999. Missing/not enough information

- 20) Targeted Population: Whether the program specifically targeted a particular minority group or sub-set of students (e.g., Aboriginal community, young mothers, low ses, etc.)

Section 5- Resources

- 21) Cost: (Excel label: "cost"): When available, information about cost is also included.
- 22) Funding Source. (EXCEL label: "Fund"): (e.g., Identify ALL funding sources listed for program: non-profit and community-based organizations, provincial or local education agencies, adult education divisions or agencies, juvenile justice agencies, private for-profit companies, public schools)
- 1. Non-profit community
 - 2. Provincial/municipal education agencies
 - 3. Juvenile justice
 - 4. Private for-profit
 - 5. Public schools
 - 999. Missing

Section 6 - Components

- 23) Program components. (EXCEL label: "Pcomp"): (e.g., academic, health and safety, life-skills, career preparation, etc Code 1=program DID involve this component and 2= did NOT involve or missing.)
- 1. Health (e.g., free lunch, medical treatment, mental health, psychologists, psychiatrists)
 - 2. Life-skills (e.g., budgeting, personal care, behaviour modification, personal mentor)
 - 3. Career preparation (e.g., career counselling, job searching, job shadowing, etc.)
 - 4. Academic support (tutoring, resource support, etc.)
 - 5. Cultural/spiritual enrichment (e.g., music, dance, religious component, etc.)
 - 6. Parental Development (e.g., family counselling, parenting classes for the participants OR for the participants' parents)
 - 7. Vocational training (e.g., certification in a trade)
 - 999. Missing
- 24) Implementation Considerations. (EXCEL label: "Impl"): Practical considerations for implementation are described in terms of support personnel required, training, additional resources, and the estimated duration of the intervention
- 25) Additional support staff (EXCEL label: "sstaff") To what extent did the program provide students with access to other adult support beside teachers (e.g., psychologists, career mentors, health care practitioners, social workers, alcohol/drug counsellors, etc.) Assistants: includes things like teaching assistants, youth care workers, business professionals, etc.)
- 1. Psychologists/counsellors
 - 2. career mentors
 - 3. health care practitioners (doctors or nurses)
 - 4. social workers
 - 5. substance abuse counsellors
 - 6. parents

- 7. peers
- 8. assistants
- 999. Missing

26) Student/teacher ratio. (EXEL label: "stratio")

27) Schedule. (EXEL label: "sched") A description of the program schedule. (e.g., interwoven within the regular classroom, flexible individualized program, student-developed timetable, etc.) (regular class time = a set timetable determined before the program started and that all students are expected to follow and participate in; flexible hours = general timetable which provides a basic structure but not specific regimen; drop-in centre = programs where there is no set schedule and students come on an as need basis; individualized= programs designed with components like the IPP (Individual Program Plan) or where each student is given an individual plan/schedule to follow and 2 students in the same program might have very different instructional schedules)

- 1. regular class time
- 2. flexible hours
- 3. drop-in centre
- 4. individualized timetable
- 999. Missing

28) Maximum number of students enrolled at one time. (EXEL label: "max") If provided enter exact number. If not reported indicate 999.

29) Culturally responsive teaching (EXEL label: "Culture") Extent to which the program celebrates the specific cultures of students, engaging in sharing about their culture and knowledge, using instructional materials that reflect cultural views and values. This is a qualitative measure; hence use your best judgment, if necessary, provide comments/references justifying why any particular assumption was made).

- 1. Yes
- 2. No
- 999. Missing

30) Individualized programs/self-expression (EXCEL label: "express"). Extent to which the program allows for personal choice and individual creativity. Examples could include allowing students to select topic for a research project, or to incorporate talents and abilities into assessment tasks. This is a qualitative measure; hence use your best judgment, if necessary, provide comments/references justifying why any particular assumption was made).

- 1. Yes
- 2. No
- 999. Missing

31) Multiple forms of assessment. (EXEL label: "assess") Extent to which the program measured multiple student measures of "success" or evaluated students using multiple kinds of assessment strategies (e.g., tests, portfolios, public presentations, etc.) This is a qualitative measure; hence use your best judgment, if necessary, provide comments/references justifying why any particular assumption was made).

- 1. Yes
- 2. No
- 999. Missing

32) Cooperative/group work (EXEL label: "groupwrk") Describes the amount of group work, cooperative learning activities built into the program design compared to other features. This is a qualitative measure; hence use your best judgment, if necessary, provide comments/references justifying why any particular assumption was made).

- 1. more time than other program features
- 2. equal time to other program features

3. less time than other program features
4. constitutes entire program
999. Not used explicitly or missing

- 33) Service learning (EXEL label: “srvlearn”) Describes the amount of service learning time that was built into the program design compared to other features. Service learning combines experiential learning and community service opportunities through things like; students build and maintain a community garden, provide free translation services for local groups or schools, remove trash from a streambed, etc. This is a qualitative measure; hence use your best judgment, if necessary, provide comments/references justifying why any particular assumption was made).
1. more time than other program features
 2. equal time to other program features
 3. less time than other program features
 4. constitutes entire program
 999. Not used explicitly or missing
- 34) Mentoring (EXEL label: “mentor”) Describes the amount of one-on-one mentorship time that was built into the program design compared to other features. Things like; job-shadowing, students being assigned an individual mentor, Big Brother Big Sister, etc. This is a qualitative measure; hence use your best judgment, if necessary, provide comments/references justifying why any particular assumption was made).
1. more time than other program features
 2. equal time to other program features
 3. less time than other program features
 4. constitutes entire program
 999. Not used explicitly or missing
- 35) Tutoring (EXEL label: “tutoring”) Describes the amount of tutoring time that was built into the program design compared to other features. Things like; peer-tutoring, afterschool homework support, summer tutoring programs, etc. This is a qualitative measure; hence use your best judgment, if necessary, provide comments/references justifying why any particular assumption was made).
1. more time than other program features
 2. equal time to other program features
 3. less time than other program features
 4. constitutes entire program
 999. Not used explicitly or missing
- 36) Use of Technology (EXEL label: “tech”) Describes the amount of computer-supported learning built into the program design. Things like computer-based tutoring systems, on-line access to course materials, designing websites, etc. This is a qualitative measure; hence use your best judgment, if necessary, provide comments/references justifying why any particular assumption was made).
1. more time than other program features
 2. equal time to other program features
 3. less time than other program features
 4. constitutes entire program
 999. Not used explicitly or missing
- 37) Career and Technical Education. (EXEL label: “vocat”) Describes the amount of career and technical education that is built into the program design. This includes, tech prep, career academies, school registered apprenticeships, student internships, career-oriented high schools, and school-based enterprises, career guidance, work-based learning, career pathways) This is a qualitative measure; hence use your best judgment, if necessary, provide comments/references justifying why any particular assumption was made).
1. more time than other program features

2. equal time to other program features
3. less time than other program features
4. constitutes entire program
999. Not used explicitly or missing

Section -7 Management

- 38) Treatment integrity. (EXCEL label: "Integ"): (Information indicating whether the treatment or intervention was implemented as specified.)
- 39) Use of an external evaluator. (EXCEL label: "Eval"): (The use of an evaluator external to program development and implementation is noted. Note: if the researcher was NOT the instructor did not deliver the intervention they can be considered and external evaluator)
 1. Yes
 2. No/Not specified
- 40) Multiple sites or studies. (EXCEL label: "Site"): (Implementation of the program in more than one setting or more than one evaluative study is noted)
 1. Yes
 2. No/Not specified
- 41) Number of years (EXCEL label: "YearRun"): the program had been operating at time of data collection.
- 42) Program type. (EXCEL label: "Ptype"): (e.g., resource rooms [separate room/teacher provides additional services], pull-out programs [within the day or even after-school]; schools-within-a-school.)
 1. Pull-out programs (during the day/after school)
 2. School within a school
 3. Alternative school
 4. Workshop or designed class
 5. Cooperative education
 6. In class
 999. Missing
- 43) Validity Rating. (EXCEL label "validity"): (e.g., *Quantitative* High=True experimental random assignment; Moderate to high=Non-equivalent pretest/posttest designs with appropriate statistical procedures and evidence of controlling for possible threats; Moderate to low =Non-equivalent pre-test/post-test design with few controls; Low:=One-group pre-test/post-test, non equivalent groups protest only, and one-group post-test only. *Qualitative* High=Data triangulation (member checks if appropriate); clearly described research procedures; personal and intellectual biases explicated stated, discussion related to negative cases or contradictory findings; Moderate to high=Detailed of descriptions of the research process; consistency between theoretical orientation and types of data collection and presentation of corroborating evidence; Moderate to low =Limited presentation/discussion of corroborating evidence; vague descriptions of research process and theoretical orientation; Low=No identifiable theoretical or methodological orientation used very limited description of research procedures, no attempts to triangulate data.
 1. High
 2. Moderate to High
 3. Moderate to low
 4. Low
 999. Missing

*Qualitative Notes: for qualitative studies make note of key “findings”, general summary of program effectiveness (positive/negative/mixed) and any particular themes identified by the author.

44.Evidence of Program Effectiveness (EXCEL label “GENOUT”) This code considers whether a program is meeting its long-term and annual performance goals. To support a “positive” or “strong positive”, the finding document must provide historical performance data showing the program’s successful progress in meeting goals. A strong positive needs to provide evidence that demonstrates a significant improvement in outcomes (statistically or qualitatively different). Evidence can also include data from performance report, a strategic plan, or other administration goals and objectives. In cases where effect sizes could not be extracted but reporting of dropout rates/graduation rates etc. are used, that data can be reflected in the code assigned here. Reports detailing learner improvements with program performance, program reports detailing rates of utilization or participation, or independent evaluations of the program’s performance may also be considered as relevant evidence (e.g., Strong positive effect reported with no overriding contrary evidence, potentially positive effects reported with no overriding contrary evidence, mixed effects or evidence of inconsistent effects, potentially negative effects with no overriding contrary evidence, strong negative effects reported with no overriding contrary evidence).

1. strong positive

2. positive

3. mixed

4. negative

strong negative

APPENDIX C- EFFECT SIZE CONVERSION FORMULAS

This appendix provides information on how the effect size estimates were determined for different research designs and statistical analysis procedures encountered during the review. According to Lipsey and Wilson (2001, p.5) the key to this kind of analysis is “defining an effect size statistic capable of representing the quantitative findings of a set of research studies in a standardized form...”

Standardized Mean Difference

This effect measure is defined as “the difference between mean outcome of the intervention group and the mean outcome of the comparison group, divided by the pooled within-group standard deviation” (What Works Clearinghouse, 2006, p1). This statistic is based on the existence of a two variable relationship in which a pre-post contrast is made on the same construct prior to exposure to the intervention and after the exposure. The difference between the two is assumed to be the result of the intervention. This pre-post contrast is compared to a control or non-intervention group. The sign of the effect size is meaningful for interpreting results. The standardized mean difference is derived by subtracting the control group pre-post difference from the treatment group pre-post difference and dividing the result by the pooled standard deviation of the measure. This results in an effect size estimate in which a positive result indicates that the treatment resulted in an increase in the outcome measure (primarily identified as graduation rate in this review) over that seen in the non-intervention group. In turn, a negative effect size estimate would indicate that the outcome variable decreased in value due to the effect of the experimental intervention.

$$d = M_1 - M_2 / \sigma_{\text{pooled}}$$

Where d is the difference between the means, $M_1 - M_2$, divided by pooled standard deviation of the two groups. Cohen’s d values were converted to Hedges g values (i.e., unbiased estimates) via equation 2 (Hedges & Olkin, 1985, p.81).

Three formulas for effect size extraction beyond the standardized mean difference were encountered in the review including: analysis of variance using F-values, t-test and proportional data in between group analysis.

Analysis of Variance using F-values

For calculation of the standardized mean difference effect size using an F-ratio (F) from a one-way ANOVA and samples sizes (n) for each group, the following formula Holmes (1980) was used. Where F is the F-ratio of the ANOVA and n_1 and n_2 are the sample sizes of the treatment and control groups, respectively.

Independent t-test

For the independent t-test (t) when either n sizes are both known or when only the total N is known) the formulas outlined by Glass et. al. (1981, pp.126-129) were used.

Proportional Data

For extracting effect sizes where probabilities of expected results can be determined through proportional data between groups, the following formula was used:

$$ES = (Pe - Pc) / \text{SQRT } (p*q)$$

Where, $p = (Pe + Pc) / 2$ & $q = 1 - p$

{Pe & Pc - probabilities of expected result (e.g., retention rate in the case of this review) for experimental and control groups respectively}

APPENDIX D- SAMPLE FORM LETTERS

(ENGLISH VERSIONS)

GOVERNMENT REPORTS/SPECIFIC REFERENCED STUDIES

Dear colleague, the Centre for the study of learning and performance at Concordia university is involved in a systematic review of the literature in the area of dropout prevention programs, especially Canadian based programs between 1990 and 2006. Its primary objective is to syntheses studies that report on the outcomes associated with these initiatives (i.e., retention rates, graduation rates, dropout rates, school retention, school -to -work transition, attitudes towards school, etc.) and provide a picture of what the research evidence suggests as “best practices”. During a web-based searches we identified a report ... (the title) (\$\$\$\$\$ INSERT URL) and think it might be considered for inclusion in our analysis. We would like to request a complete copy of this report for consideration. You can forward a copy via e-mail or via regular mail to the address listed above. Please contact me with questions or comments.

Thank you for your assistance.

(FRENCH VERSIONS)

DIRECTIONS D'ÉCOLE/ADMINISTRATIONS DE COMMISSION SCOLAIRE

Madame,

Monsieur,

Le Centre d'études sur l'apprentissage et la performance de l'Université Concordia mène actuellement un recensement méthodique de la littérature portant sur les programmes de prévention de l'abandon scolaire, et plus particulièrement sur les programmes mis en oeuvre au Canada entre 1990 et 2006. Notre objectif principal consiste à faire la synthèse des études qui présentent des résultats à cet égard (taux de persévérance scolaire, d'obtention du diplôme et d'abandon scolaire, transition entre l'école et le travail, attitudes envers les études, etc.) et à dégager ce que la littérature étudiée désigne comme des « pratiques exemplaires ». Le ministère de l'Éducation de votre province a désigné votre programme ... (titre) comme initiative de prévention de l'abandon scolaire, et nous pensons qu'il pourrait être utile à notre analyse. Serait-il possible d'obtenir un exemplaire intégral de tout rapport ou document relatif à ce programme? Vous pouvez nous le faire parvenir par courriel ou par la poste à l'adresse indiquée ci-dessus. N'hésitez pas à communiquer avec nous si vous avez des questions ou des commentaires à ce sujet.

Nous vous remercions à l'avance de votre aide.

APPENDIX E- SUMMARY OF EFFECT SIZE EXTRACTIONS

<i>Author</i>	<i>Year</i>	<i>Outcome Type¹</i>	<i>ES (Hedges G)</i>
Vitaro *	1999	Social/Behavioural	0.36
Saint-Laurent*	1996	Psychological	0.33
Saint-Laurent	1996	Psychological	0.08
Saint-Laurent	1996	Academic/Cognitive	-0.01
Saint-Laurent	1996	Academic/Cognitive	0.10
Theoret*	2000	Psychological	0.39
Theoret	2000	Physical Presence	0.21
Remizoff*	1990	Psychological	0.81
Remizoff	1990	Psychological	0.50
Fortin*	1998	Social/Behavioural	0.05
Fortin	1998	Social/Behavioural	0.06
Fortin	1998	Social/Behavioural	0.05
Fortin	1998	Physical Presence	-0.02
Fortin	1998	Physical Presence	-0.47
Breckon*	1996	Physical Presence	0.45
Locke*	1993	Physical Presence	0.16
Locke	1993	Physical Presence	0.41
Locke	1993	Physical Presence	0.23
Mutadi*	1990	Physical Presence	1.06
Mutadi	1990	Physical Presence	0.72
Mutadi	1990	Physical Presence	1.06
Mutadi	1990	Psychological	0.87
Mutadi	1990	Psychological	0.73
Mutadi	1990	Psychological	1.80
Hahn	1994	Physical Presence	0.82
Hahn	1995	Physical Presence	0.56

Hahn	1996	Physical Presence	0.41
Hahn	1997	Physical Presence	0.39
Stadford	1993	Physical Presence	0.27
McHale	1994	Physical Presence	-0.18
Hayward	1995	Physical Presence	0.03
Hayward	1995	Physical Presence	0.08
Hayward	1995	Physical Presence	0.65
Hayward	1995	Physical Presence	0.18
Hayward	1995	Physical Presence	0.27
Hayward	1995	Physical Presence	0.47
Hayward	1995	Physical Presence	0.28
Hayward	1995	Physical Presence	-0.15
Hayward	1995	Physical Presence	0.02
Hayward	1995	Physical Presence	0.00
Hayward	1995	Physical Presence	-0.31
Hayward	1995	Physical Presence	0.16
Hayward	1995	Physical Presence	0.39
Hayward	1995	Physical Presence	0.40
Hayward	1995	Physical Presence	-0.24
Hayward	1995	Physical Presence	-0.67
Hayward	1995	Physical Presence	0.81
Hayward	1995	Physical Presence	-0.40
Hayward	1995	Physical Presence	0.18
Hayward	1995	Physical Presence	0.28
Hayward	1995	Physical Presence	0.10
Hayward	1995	Physical Presence	0.25
Thorton	1995	Physical Presence	0.06
Elliot	2001	Physical Presence	0.49
Elliot	2001	Physical Presence	0.30

Munoz	2002	Physical Presence	0.39
Trautman	2003	Physical Presence	0.13
Counts	1992	Academic/Cognitive	1.07
Carruthers	1999	Physical Presence	0.34
Carruthers	1999	Psychological	0.62
Sinclair	2005	Physical Presence	0.38
Snipes	2005	Physical Presence	0.11
Snipes	2005	Physical Presence	-0.26
Snipes	2005	Physical Presence	0.01
McPartland	1991	Physical Presence	0.18
McPartland	1991	Academic/Cognitive	-0.10

APPENDIX F- SUMMARY OF EFFECT SIZES BY COUNTRY OF STUDY ORIGIN

ES from Canadian Based Programs

Saint-Laurent	1996	Academic/Cognitive	-0.01
Saint-Laurent	1996	Academic/Cognitive	0.10
Fortin	1998	Physical Presence	-0.47
Fortin	1998	Physical Presence	-0.02
Locke	1993	Physical Presence	0.16
Theoret	2000	Physical Presence	0.21
Locke	1993	Physical Presence	0.23
Locke	1993	Physical Presence	0.41
Breckon	1996	Physical Presence	0.45
Mutadi	1990	Physical Presence	0.72
Mutadi	1990	Physical Presence	1.06
Mutadi	1990	Physical Presence	1.06
Saint-Laurent	1996	Psychological	0.08
Saint-Laurent	1996	Psychological	0.33
Theoret	2000	Psychological	0.39
Remizoff	1990	Psychological	0.50
Mutadi	1990	Psychological	0.73
Remizoff	1990	Psychological	0.81
Mutadi	1990	Psychological	0.87
Mutadi	1990	Psychological	1.80
Fortin	1998	Social/Behavioural	0.05
Fortin	1998	Social/Behavioural	0.05
Fortin	1998	Social/Behavioural	0.06
Vitaro	1999	Social/Behavioural	0.36

ES from Non-Canadian Based Programs

McPartland	1991	Academic/Cognitive	-0.10
------------	------	--------------------	-------

Counts	1992	Academic/Cognitive	1.07
Hayward	1995	Physical Presence	-0.67
Hayward	1995	Physical Presence	-0.40
Hayward	1995	Physical Presence	-0.31
Snipes	2005	Physical Presence	-0.26
Hayward	1995	Physical Presence	-0.24
McHale	1994	Physical Presence	-0.18
Hayward	1995	Physical Presence	-0.15
Hayward	1995	Physical Presence	0.00
Snipes	2005	Physical Presence	0.01
Hayward	1995	Physical Presence	0.02
Hayward	1995	Physical Presence	0.03
Thorton	1995	Physical Presence	0.06
Hayward	1995	Physical Presence	0.08
Hayward	1995	Physical Presence	0.10
Snipes	2005	Physical Presence	0.11
Trautman	2003	Physical Presence	0.13
Hayward	1995	Physical Presence	0.16
McPartland	1991	Physical Presence	0.18
Hayward	1995	Physical Presence	0.18
Hayward	1995	Physical Presence	0.18
Hayward	1995	Physical Presence	0.25
Hayward	1995	Physical Presence	0.27
Stadford	1993	Physical Presence	0.27
Hayward	1995	Physical Presence	0.28
Hayward	1995	Physical Presence	0.28
Elliot	2002	Physical Presence	0.30
Carruthers	1999	Physical Presence	0.34
Sinclair	2005	Physical Presence	0.38

Hayward	1995	Physical Presence	0.39
Munoz	2002	Physical Presence	0.39
Hahn	1997	Physical Presence	0.39
Hayward	1995	Physical Presence	0.40
Hahn	1996	Physical Presence	0.41
Hayward	1995	Physical Presence	0.47
Elliot	2001	Physical Presence	0.49
Hahn	1995	Physical Presence	0.56
Hayward	1995	Physical Presence	0.65
Hayward	1995	Physical Presence	0.81
Hahn	1994	Physical Presence	0.82
Carruthers	1999	Psychological	0.62

APPENDIX G- LIST OF INCLUDED STUDIES/REPORTS

Canadian Studies

- 1.Evaluation of Alberta start projects: Provincial perspective. Final report.68.
- 2.Stay in--you win: Module Four. Project overview.86.
- 3.Alpern, Michael. The integrated occupational program: Building bridges for at-risk students.1991; 20.
- 4.Assogba, Yao. Raccrocher l'autre jeunesse: Une expérience communautaire de raccrochage scolaire dans l'Outaouais. *Apprentissage Et Socialisation*. 1992; 15(1):39-48.
- 5.Bell, David and Society for the Advancement of Excellence in Education. *Sharing our success: Ten case studies in aboriginal schooling*.
- 6.Bouchard, P. (1994). Le projet Foyer: une mesure d'aide aux décrocheurs du premier cycle du secondaire. *Vie Pédagogique* 91, 12-14, 39.
- 7.Bouchard, Pierrette and St-Amant, Jean-Claude. Le retour aux études: Les facteurs de réussite dans quatre écoles spécialisées au Québec. *Revue Canadienne De L'Éducation*. 1996; 21(1):1-17gagne
- 8.Breckon, Dave. *Beyond school: A pilot project*. in: Witt, P. A. and Crompton, J. L., Eds. *Recreation programs that work for at-risk youth: The challenge of shaping the future*. State College, PA: Venture Publishing, Inc.
- 9.Brown, Robert S. Challenges and potential of mentoring at-risk students: A literature review. *ERS Spectrum*. 1996,
- 10.Brown, R. S. Tracking of grade 9 and 10 students in the Toronto board. September 1991-September 1995.15; ISSN: 0316-8786.
- 11.A two-year evaluation of the change your future program at the Toronto board of education. 1991-1993. (208):26; ISSN: 0316-8786.
- 12.Brigham, S; Taylor, A. (2006). Youth apprenticeship programs for aboriginal youth in Canada: Smoothing the path from school to work? *Canadian Journal of Native Education*, 29,2 165-182.
- 13.Campbell, Donald S. Evaluating an intervention for at-risk adolescents: The case of the breakaway company. *Canadian Journal of Education*. 1995, Fall ; 20(4):437-454 ; ISSN: 0380-2361.
- 14.Chinien, Chris A. and Boutin, France. Qualitative assessment of cognitive-based dropout prevention strategy. *The High School Journal*. 2002; 85(2):1-11.
- 15.Fortin, L.; Picard, Y. (1998). Les effets d'un programme d'intervention multidimensionnel auprès d'élèves à risqué de décrochage scolaire au secondaire. *Revue Quebecoise de psychologie* 19, 2, 125-145.
- 16.Gagne, A. (1996). Success at Contact: the argument for alternative school for at-risk youth. *The Alberta Journal of Educational research*. XLII , 3, 306-324.
- 17.Housego, Billie E. J. (1999). "Outreach Schools: An Educational Innovation". *Alberta Journal of Educational Research*, XLV(1), pp. 85-101
- 18.Kearns, John R. and others. The London project: The implementation of invitational strategies at the secondary school level.1990; 45.
- 19.Kelly, Deirdre M. Dilemmas of difference: Van tech's schools-within-a-school model. *Alberta Journal of Educational Research*. 1996, Sep; 42(3):293-305; ISSN: 0002-4805.
- 20.Langevin, L. (1995). Collaboration entre l'école et l'université: l'évaluation des effets d'un service de soutien pour décrocheurs potentiels. *Vie Pédagogique* 94, 4-5.
- 21.Locke, D. The longitudinal effects of a wilderness adventure program on potential dropouts. (1993). An unpublished masters thesis. Department of Educational Psychology Memorial University of
- 22.McCluskey, Ken W.; Baker, Philip A.; Bergsgaard, Mike, and McCluskey, Andrea L. A. (The University of Winnipeg). *Creative problem solving in the trenches: Interventions with at-risk populations*. Creative Problem Solving Group: Williamsville, NY; 2001, February; 308.

23. McCluskey, Ken W.; Baker, Philip A., O'Hagan, S.C., Treffinger, D. J. (1998). Recapturing at-risk, talented high school dropouts: A summary of the three year Lost Prizes project. *Gifted and Talented International* 13, 73-78.
24. Miclot, M. (2000). La voie technologique: une solution au décrochage. *Vie Pédagogique*, 115, 42-43.
25. Mutadi, Neal A. Common behaviors and characteristics of high school dropouts and the effect of support strategies on dropout rates. *Dissertation Abstracts International*. 1991; 52(3):765.
26. Remizoff, Mary. Evaluation of the program for parenting adolescents. 256.
27. Rowan; Norman, and The Pathways to Education Program TM (find full author names and more details). Pathways [Web Page]. 2006; Accessed 2007 Dec 28. Available at: www.pathwaystoeducation.ca.
28. Saint Laurent, Lise. Pier: an inclusive model for at-risk students. 1996; 27.
29. Soudack, Avi. Research evaluation of the Toronto board of education community-based education for work, career and life (CWCL) pilot project, 1993-94. *Guidance and Counselling*. 1996, Spring; 11(3):31-42 ; ISSN: 0831-5493.
30. Théorêt, Manon; Garon, Roseline, and Himech, Mohamed. évaluation d'une intervention de mentorat visant à réduire le risque d'abandon scolaire. *Revue Canadienne De Psycho-Education*. 2000; 29(1):65-86.
31. Theoret, M. L'importance d'un mentor pour les élèves qui risquent d'abandonner l'école. (1996). *Vie Pédagogique* 98, 42-44.
32. Brossard, L. (1996). Un meilleur accès au savoir pour tous les élèves. *Vie Pédagogique* 97 37-42.
33. Keeping Kids in School (2005). The new school Canada: A Comprehensive approach to building literacy skills, academic knowledge, and the behaviours required for success in school and in life. Literacy BC.
34. Bajnok, B.; Eaton, D. (1993). Stay in school initiative Bishop Carroll style. *Challenge in Educational Administration* 30, 3, 10-12.
35. Vitaro, Frank; Brendgen, Mara, and Tremblay, Richard E. Prevention of school dropout through the reduction of disruptive behaviors and school failure in elementary school. *Journal of School Psychology*. 1999; 37(2):205-226.
36. Volpe, Richard; Clancy, Christine; Buteau, Cindy, and Tilleczeck, Kate. Effective Ontario initiatives to retain secondary students at risk of dropping out of school. Toronto, Ont.: University of Toronto; 1998.
37. Yeo E. (1990). A follow up study of 1988 summer program for potential early school leavers in the province of New Brunswick. Unpublished master' thesis. University of Prince Edward Island, 1990.
38. The stay in school initiative: Ideas that work (1996). Human resources Development Canada

Non-Canadian Studies

1. Allen, Joseph P.; Philliber, Susan; Herrling, Scott, and Kuperminc, Gabriel P. Preventing teen pregnancy and academic failure: Experimental evaluation of a developmentally based approach. *Child Development*. 1997, August; 68(4):729-742.
2. Austin Independent School District, TX. Office of Program Evaluation. Gear up Austin: Impacting lives project, 2000-2001. 2001, December; AISD-Pub-00.19. 81 p
3. Carruthers, William L. Collected evaluations on the WCPSS alternative school program for students with long-term suspensions. Executive summary. Eye on evaluation. 1999, July; E&R Report no. 00.01. 12 p.
4. Casey, Arthur C.; McSwain, James, and Beach, Don M. The cooperative alternative program high school: A dropout prevention plan. *NASSP Bulletin*. 1993; 77:105.
5. Counts, E. (1992). The impact of participation in an "extended day" program upon academic achievement. Paper presented at the annual meeting of the Mid_south Educational Research Association. Knoxville, TN.

6. Elliott, Marc N.; Hanser, Lawrence M., and Gilroy, Curtis L. Evidence of positive student outcomes in JROTC career academies. Prepared for the Office of the Secretary of Defence.: National Defence Research Institute.
7. Faircloth, Beverly S. and O'Sullivan, Rita. Strategy 17: NovaNET online learning system year 2 program evaluation. Wake County Public School system safe schools grant. Chapel Hill, SC: School of Education, North Carolina University, Chapel Hill; 2001, December 1231 p.
8. Galleta-Bruno, Diana. A comprehensive dropout-prevention program to increase the number of Spanish-speaking ESL students remaining in high school: Nova Southeastern University; 1995Ed. D. Practicum.
9. George State Department of Education, Atlanta. Evaluation of the Georgia crossroads alternative education program, July 1997-June 2000. Volumes I and II. 2000, October; 227 p.
10. Groth, Cori. Dumping ground or effective alternative: Dropout-prevention programs in urban schools. *Urban Education*. 1998; 33:218.
11. Hahn, Andrew and others. Evaluation of the quantum opportunities program (QOP). Did the program work? A report on the post secondary outcomes and cost-effectiveness of the QOP program (1989-1993). Waltham, MA: Center for Human Resources, Brandeis University; 1994, June; 78 p.
12. Hayward, Becky Jon and Tallmadge, G. Kasten (American Institutes for Research in the Behavioral Sciences, Washington D.C.). Strategies for keeping kids in school: Evaluation of dropout prevention and re-entry projects in vocational education. Final report. 1995, June; 269 p.
13. Johns, Susan; Kilpatrick, Sue; Mulford, Bill, and Falk, Ian. Thinking outside the box: A remote VET in schools program challenges traditional boundaries. 2001, Research to Reality: Putting VET Research to Work. Proceedings of the Australian Vocational Education and Training Research Association (AVETRA) Conference (4th, Adelaide, Australia, March 28-30, 2001).
14. Karl, Jane Wood and Karl, Luis C. An experimental design survey of attitude changes towards education by at-risk students placed in an alternative high school program. Rhinelander, WI: Nicolet Area Technical College; 2000, Feb60 p. Practicum paper.
15. Lever, Nancy; Sander, Mark A.; Lombardo, Sylvie; Randall, Camille; Axelrod, Jennifer; Rubenstein, Michelle, and Weist, Mark. A dropout prevention program for high-risk inner-city youth. *Behaviour Modification*. 2004; 28:513.
16. Madison, Marion; Marson, Alvin, and Reese, Kiki. Passageway: An avenue into the future In: Rural special education for the new millennium. Conference proceedings of the American Council on Rural Special Education (ACRES) (19th, Albuquerque, New Mexico, March 25-27, 1999).: 1999, March5 p.
17. Matos, Carmen M. Decreasing early school withdrawal and retention of Hispanic and other limited English proficient (LEP) students through counselling and arts/career exploration.: Nova Southeastern University; 1998, September 3085 p.Ed. D. Practicum.
18. McHale, Sharon. Evaluation of the career education center (1993-1994). Brooklyn, NY.: Office of Educational Research, New York City Board of Education; 1994, August 01; 70 p.
19. McPartland, James M. and Nettles, Sandra Murray. Using community adults' advocates or mentors for at-risk middle school students: A two-year evaluation of project RAISE. *American Journal of Education*. 1991, August; 99(4):568-586.
20. Mitchell, Stephanie and Waiwaiiole, Gayle. Interim evaluation of in-district alternative education high school programs. Research and Development Department of Portland High Schools (?); 2003, February.
21. Munoz, Marco A. Facing the challenges of at-risk students in urban school districts: The impact of an attendance and dropout prevention program in a non-traditional school. 200220 p.
22. Opuni, Kwame A. and others. Beating the odds summer school: A dropout prevention program for at-risk students: Department of Research and Evaluation of the Houston Independent School District (?); 1990; 19 p.
23. O'Sullivan, Rita G. Model dropout prevention program at Reidsville Middle School: A case study evaluation. Paper presented at the Rural Education Symposium of the American

- Council on Rural Special Education and the National Rural Small Schools Consortium (Tucson, AZ, March 18-22, 1990).: 1990, March; 12 p.
24. Rumberger, Russell W. and Larson, Katherine A. Keeping high-risk Chicano students in school: Lessons from a Los Angeles middle school dropout prevention program. Paper commissioned for At-Risk Evaluation.: 1992, September; 56 p.
 25. Sinclair, M.; Christenson, S.; Thurlow, M. (2005). Promoting school completion of urban secondary youth with emotional and behavioral disabilities. *Exceptional children* 71,4, pp.465-482.
 26. Snipes, J.; Holton, G.; Doolittle, F.; Szejnberg, L. (2006). Striving for student success: the effect of project GRAD on high school student outcomes in three urban school districts. Report available on WWW: www.mdrc.org.
 27. Stradford, Charmaine W. Implementation of a rural program to reduce the dropout rate of 9th- and 10th-grade at-risk students.: Nova University; 1993, June 135 p. Ed. D. Practicum.
 28. Thorton, H. (1995). Staying in school: A technical report of three dropout prevention projects for middle school students with learning and emotional disabilities. Santa Barbara: California University.
 29. Trautman, T. (2001). Credit recovery: A technology-based intervention for dropout prevention at Wichita Falls High School. American Education Corporation.
 30. Wendt-Kenswick, Karen and Paredes, Theresa. Delta program third year implementation. Feedback.: Office of Program Evaluation of the Austin Independent School District (?); 1998, September; Publication No. 97.14. 9 p.