Supporting Students: A GRADE Analysis of the Research on Student Wellness and Classroom Mental Health Support

Susan Rodger¹, Renelle Bourdage¹, Kaitlin Hancock¹, Rebecca Hsiang¹, Robyn Masters¹, and Alan Leschied¹

Abstract
Recommendations derived from research evidence regarding program implementation in school-based mental health [SBMH] require knowledge of the intervention outcomes as well as the potential to translate program components into schools. The Grading of Recommendations, Assessment, Development, and Evaluation (GRADE) analysis was applied to major areas of the SBMH literature in addressing the areas of effectiveness and the strength of recommendation in implementation. Nine areas were addressed: emotional intelligence, stigma reduction, mindfulness, anxiety, depression, addictions, suicide prevention, trauma, and eating disorders. Ninety-eight studies were retrieved across the nine areas. Effect sizes based on reported outcomes and estimates on the strength of recommendation were generated in each of the nine areas of interest. These results provide an overview of the quality of the evidence that will be of relevance to school personal in making program selections.

Keywords
GRADE analysis, school intervention, effectiveness

Introduction
Concern for child and youth mental health cannot be over stated. Most major mental health disorders have their onset prior to the age of 25 years. Recent data in Canada indicate that up to 25% of Canadians up to the age of 19 years experience at least one

¹The University of Western Ontario, London, Canada

Corresponding Author:
Alan Leschied, Althouse College, The University of Western Ontario, 1137 Western Road, London, Ontario, Canada N6A 3K7.
Email: leschied@uwo.ca
mental health problem. These include in descending order of representation: an eating disorder (ED), anxiety, attention deficit/hyperactivity disorder (ADHD), mood disorders, psychosis, and substance abuse. Suicide accounted for 20% of all deaths in young people aged 10 to 29 years in 2013 (Butler & Pang, 2014). These data stand in stark contrast to the fact that most youth, even after they have been assessed and receive a mental health diagnosis, are unable to access necessary and timely mental health intervention. The shortage of child and youth mental health resources is the most frequently cited reason for the inability to meet this need.

Momentum is now firmly established regarding the evolving role of schools in the delivery of mental health services for children and youth as an innovative way to reach the number of young people who are now reporting distressing levels of a mental health disorder. The well-known statistic of one in five of our young people experiencing clinical levels of distress belies the fact that many more children and youth, while not reaching symptom severity levels meeting diagnosable criteria, are also in need of some form of counseling or clinical support (Flett & Hewitt, 2013). Many of the school’s efforts target universal prevention in mental health literacy, curriculum-based delivery of specific strategies, as well as coordinated efforts for youth whose symptom severity requires more intensive and coordinated services (Kutcher & McLuckie, 2009).

There is now mounting evidence in Canada to support school-based mental health (SBMH) as reflected in the peer-reviewed literature (Leschied, Saklofske, & Flett, 2013), in books (Weist, Lever, Bradshaw, & Owens, 2014) along with broad scale reviews as provided through the Mental Health Commission of Canada (2013a). The momentum with regard to knowledge that can guide schools in selecting those programs that can best fit within the current emphasis on SBMH appears to be firmly planted.

**Evidence for Situating Schools on the Front Line in Providing Mental Health Services**

There is now broad-based support for schools to enact policies and programs that support the delivery of SBMH services, either as adjunct resources to schools or as part of the school curriculum. In the United States, The Every Student Succeeds Act (ESSA) supports an emphasis on Social and Emotional Learning that now appears in the new Federal Education Law (retrieved from the Collaborative for Academic and Social Emotional Learning [CASEL] website: http://www.casel.org/federal-policy-and-legislation/)

The Mental Health Commission of Canada (2012) in their report *Changing Directions Changing Lives* called on schools to increase their capacity to provide mental health services to young people. At the provincial level, there is clear evidence for support of the federal priority in viewing schools as places where mental health promotion can be actively promoted. The province of Ontario, for example, recently established School Mental Health Assist, a provincewide mental health initiative to guide the implementation of SBMH strategies with the support from a designated
mental health champion in each of the 31 English Public, 29 English Catholic, four French Public, and eight French Catholic School Boards (Centre for Addiction and Mental Health, 2016). The high-profile Ontario education advocacy group People for Education (2013) recently disseminated a discussion paper promoting the view that social emotional learning and mental health status in schools should be equated with, and measured against, school outcomes along with the more traditional areas of numeracy and literacy.

Impact of the Emphasis on Mental Health on the School System

This emphasis on SBMH comes at a time when there are numerous other competing priorities for the education system. There is the perennial concern for a lack of skill building in core areas such as science and mathematics education and concern for school safety reflected in the promotion of safe school policies and practices. However, with the publication of numerous reviews underscoring the interrelationship between traditional school learning outcomes and social emotional adjustment produced by Kimberley Hoagwood and Johnson (2003) and Hoagwood et al. (2007), there is an increasing appreciation for viewing what Suldo, Gormley, DuPaul, and Anderson-Butcher (2014) indicate is the bidirectional nature of personal adjustment and academic achievement.

Student well-being with an emphasis on SBMH is coming at a time of increasing appreciation for the fact that educators are reporting high rates of occupational stress themselves (Rodger & Marko, 2015). This is a significant issue, as Arens and Morin (2016) reported a direct negative association between teachers’ emotional exhaustion and student’s academic achievement reflected in grades, standardized achievement scores, and school satisfaction. If we are to promote SBMH, it needs to be accomplished mindful of the needs and well-being of teachers. There is work being done to address this area, most notably in efforts such as by Bartel, Weston, et al. (2005) in providing programs and support for teachers acknowledging the degree of stress they experience by virtue of being in a high stress profession. What is in our view incumbent on advocates within the SBMH movement is to provide support for teachers that facilitates access to current information and resources that enable teachers to make effective decisions regarding their response to students who they view as in distress. The purpose of the current review is to report on programs at the school and classroom level regarding the strength and quality of the evidence for effective strategies targeting general student well-being along with specific mental health issues.

Selected Outcome Studies Relate to the Delivery of Mental Health Services for Teachers and Students Within Schools

The principle of effective strategies targeting general student well-being draws on the assumption that there currently exist a substantive number of outcome studies examining emotional competency and various mental health interventions that can be
delivered within schools. In part, momentum in SBMH not only recognizes that schools are among the best places in which to situate child and youth wellness and mental health programs, but in addition, classroom teachers are de facto on the front line of the mental health service delivery system, with classroom teachers often being the most likely person to identify a young person in distress.

The authors of this review are well acquainted with both teacher professional development and teacher preparation programs that emphasize pedagogical preparation to the exclusion of student emotional wellness (Rodger & Marko, 2015). However, we acknowledge the systemic reality that a mental health system for children and youth can never adequately meet the needs of children and youth without including the critically and vitally important role of teachers (Climie, 2015; Tolan & Dodge, 2005). In addition, it has become all too clear that the mental health status of students has a significant impact on their learning potential and academic outcomes and as a result requires educators to be aware of such programs (Durlak, Weissberg, Dymnicki, Taylor, & Schellenger, 2011).

**Targeted Areas of Service are Clinically Relevant**

The seminal study by Dan Offord and his colleagues published almost 40 years ago (Offord et al., 1987) indicated that one in five young people experience a level of distress potentially leading to a diagnosable mental health disorder. However, while this one in five statistic is replicated across cultures in Australia (Sawyer et al., 2001) and the United Kingdom (Ford, Goodman, & Meltzer, 2003) what is not reflected in epidemiological studies is the large number of young people who experience a level of distress that, while not reaching a diagnosable level of a disorder, remains a critical concern. Flett and Hewitt (2013) referred to this group of young people as “flying under the radar” of detection in masking their distress in a myriad of ways. Teachers may see students grieving the loss of a grandparent, or experiencing the distress from online bullying. Although the grief, sadness, or anger may be transient in nature, it can interfere with the student’s academic focus or peer functioning. Emotional distress is not confined solely to those young people who may experience severe depression or debilitating anxiety but also relates to those whose sadness or social withdrawal is an indication of a transient, albeit low-grade disorder that, with the right amount of support and understanding, can prevent more serious disorders at a later time. Hence, the choice of studies that were reviewed reflected the broad range of disorders and programs from prevention (i.e., addictions) to a response to an existing disorder (i.e., anxiety or trauma resulting from bullying).

Beyond the severity of disorders that are addressed in classroom-based interventions is the nature of the disorders themselves. Here, the selection of interventions ranged from those that were universal, including emotional intelligence (EI), stigma reduction (SR), and mindfulness, to the targeted service areas that are frequently identified as the basis of greatest concern by educators (Kutcher, 2013), namely, anxiety, depression, addictions, suicide prevention, trauma, and EDs. Assignment of each study to the targeted clinical area was made following a review of the program’s primary intended intervention.
Targeted Areas of Service are Developmentally Appropriate

The selection of studies was mindful of being developmentally sensitive across both elementary and secondary school age youth. Developmental psychopathology has taught us that while certain disorders such as depression may be evidenced by similar symptoms across the life span (i.e., sleep disturbance, sadness and loss of hope), the age or developmental stage at which such symptoms are experienced will be manifest in very unique ways. Hence, the studies that were chosen and reported on reflected the context of age or grade in which they were delivered.

All Studies Were Based on Group Comparisons

The final criterion in the selection of studies for the Grading of Recommendations, Assessment, Development, and Evaluation (GRADE) analysis involved group comparisons. To be included in the analysis, the methodology had to reflect a comparison of the treated group with at least one other contrast group. The contrast group could include an untreated, usual service, or delayed treatment condition and the outcome had to reflect the benefits of intervention through self report, teacher ratings, and behavioural or observational ratings.

Method

To be most informative at this juncture in the evolution of SBMH, the authors drew on the GRADE methodology in providing a review of best practice approaches in SBMH (Ansari, Tsertsvadze, & Moher, 2009; Guyatt et al., 2008). GRADE, while similar to meta-analysis in examining the strength of evidence in a given area of intervention, also provides a systematic appraisal of the quality of the evidence (high, moderate, low, and very low quality) and strength of recommendations defined, according to Terracciano, Brozek, Compalati, and Schünemann (2010), as reflecting confidence in the achieved outcome. Recommendations are then made regarding best practice in specifically targeted areas within SBMH interventions. It is relevant to note that when combining the measurement of strength and quality of evidence using a GRADE analysis, caution is necessary when a strength rating may be emphasized in favor of the higher quality rating, despite that the latter may be more informative in terms of determining whether to adopt the program in question.

Principles Guiding the Evidence in Program Selection

There were four principles that guided the selection of studies within the targeted areas of intervention for SBMH: (a) Selected outcome studies were related to service delivery within schools either schoolwide based within classrooms, (b) targeted areas of service were relevant for teachers who have identified certain issues in their students, (c) targeted areas of service were developmentally appropriate and extended from elementary to secondary school, and (d) all outcome studies were based on group
comparisons with methodologies that attempted to control for extraneous variation reflected in randomized clinical trials (RCTs) or quasi experimental designs.

Procedure

Literature was retrieved through the most common indexed peer reviewed sources including Google Scholar, PsycINFO, and ProQuest. Search terms were selected that focused on school-based and classroom-focused delivery of programs that reflected the nine areas identified as being of most relevance to teachers, namely, EI, SR, mindfulness, anxiety, depression, addictions, suicide prevention, trauma, and EDs. Abstracts of all initially selected studies were reviewed to insure that the focus of the programs, the context of delivery, nature of design, and outcome measures matched the search criteria and were reflected in the study.

Criteria for inclusion. The inclusion criteria in the selection of studies beyond the areas of interest included two group designs with at least one measured outcome period. For a study to be included in the GRADE analysis, a publication date within the last 10 years was required (June 2005 to June 2015). The only exception to the publication date restriction was in the suicide program area. This area required retrieval dating 20 years, as the number of comparison studies examining school-based suicide prevention efforts has not developed in recent years, yet interest and concern remains high. These searches within the inclusion criteria resulted in 98 studies in the analysis.

Judging the quality of the evidence. Study design in the GRADE analysis is critical in judging the quality of the evidence. All relevant RCTs and studies that compared a treated group with a contrast group were included in the analysis. Outcome measures that were accepted for inclusion were student self report of well-being, teacher observation, and measures of behavioural and emotional functioning.

Quality assessment. Effect sizes were calculated based on means and standard deviations, t tests, or F tests. The Effect Size Generator was utilized for studies that used t tests, F tests, or regression (Becker, 1998). Cohen’s d was used for comparing treatment outcomes (Dunst, Hamby, & Trivette, 2004; Thalheimer & Cook, 2002). To determine the average effect size for each outcome, the effect sizes of the relevant studies were averaged. Determination of the strength of the effect size followed from Cohen (1988) who determined that a small ES = 0.20, medium ES = 0.50, and a large ES = 0.80. It is important to note, however, that in using Cohen’s d, the application of the strength (small, medium, large) is dependent on the specific area of study, in the current case two or more group designs comparing school-based interventions. Cohen originally noted that even a small effect size can be theoretically and/or practically meaningful. The current study applied Cohen’s cutoffs. Readers may want to look more specifically at which area is being reported on in the context of what is known in that area in assessing the relevance of these particular cutoff scores.
Calculation of effect sizes. Following previous procedures (S. Stewart, Leschied, den Dunnen, Zalmanowitz, & Baiden, 2013; Van Adel, Geier, Perry, & Reitzel, 2011), calculation of effect sizes for the GRADE analysis was based on means and standard deviations, proportions benefiting from an intervention, or values of chi-square analyses, t tests, or F tests. The Effect Size Generator was utilized for studies that used t tests, F tests, or regression (Devilly, 2004). For categorical outcomes, Cohen’s d was calculated using G*Power with chi-square values and sample size (Buchner, Erdfelder, & Faul, 1996). In certain circumstances, calculation of Cohen’s d required assumptions about the size of the sample for the analysis based on the sample description rather than the number of responses per calculated calculation (Lee, Bright, Svoboda, Fakunmoju, & Barth, 2011). In determining the average effect size for each outcome variable within each intervention category, the effect sizes of the relevant outcomes were averaged such that there was only one effect size for each outcome variable.

GRADE coding. GRADE coding involved researchers reviewing each study in the context of the ES given the outcomes, the quality of evidence as reflected in the methodology, and the strength of recommendation that reflected the advantages of the intervention. Quality was graded on a 4-point scale that reflected the nature of the research design (1 = very low, 2 = low, 3 = moderate, and 4 = high) with strength given a rating of either 1 = weak or 2 = strong (Van Adel et al., 2011). The factors that lead to downgrading the quality of studies included modification of the RCT (e.g., matching), lack of randomization, lack of fidelity measures, and smaller sample sizes. Studies were given a weak recommendation if they yielded small effect sizes, had a high attrition rate, or did not seem practical for implementation in a school setting reflected in a program’s high cost, length of intervention and/or training, and not being manualized.

For this review, two of the authors of this article independently graded different topics within the 98 articles. A third independent rater coded 10 articles to assess the level of agreement between the two raters (five from each of the primary two raters) that generated an agreement of 95% on the measure of the quality of the evidence. With the 95% agreement, the original rating of each rater was used for the final rating.

Results

The results summarize the GRADE ratings for each program in each of the nine areas. The first three areas relate to general coping, including EI, SR, and mindfulness, followed by the six targeted mental health areas of depression, addictions, anxiety, trauma, suicide prevention, and EDs.

EI

Broadly speaking, EI reflects the emotional competencies that identify, process, and regulate emotions (Austin, Saklofske, & Egan, 2005). EI is highly relevant in supporting children and youth in the developmental process of self-regulation and emotional awareness of the self and others (Mayer, Salovey, & Caruso, 2004). Educators have seized on
EI as a relevant construct in creating school cultures that are supportive, with advocates suggesting that higher EI scores promoting social emotional learning are associated with improved academic success (Elias & Arnold, 2006). As a result, a number of strategies have developed that are systemwide and schoolwide as well as classroom specific.

The quality of the evidence for EI was comprised of nine RCTs and one quasi-experimental study and was rated from low to high. The sample sizes ranged from 510 to 7,846 participants. The mean age of participants for all 10 EI programs was 10.2 years ($SD = 2.95$). The average program duration was 30.3 weeks ($SD = 12.97$). Studies were given a weak recommendation if they yielded small effect sizes, had a high attrition rate, or did not seem practical for implementation in an ordinary school setting (e.g., high cost, lengthy intervention and/or training, not manualized).

All interventions for EI were considered lengthy, averaging slightly more than 30 weeks; the shortest intervention, The Gatehouse Project (Bond et al., 2004), took place over a 10-week period. Some of the EI interventions were implemented over multiple school years. The frequency of all interventions was quite variable; however, most interventions took place once or twice a week, with lessons lasting 40 min to 1 hr. All interventions were manualized, involving scripts and lesson plans to increase transferability or ease of implementation. Finally, most programs (70%) involved more intensive training requiring attendance at a workshop of more than 2 days along with some form of ongoing supervision during implementation. Although this negatively affected their transferability, this level of supervision seemed necessary for successful implementation. All of these interventions were integrated within the everyday curriculum (i.e., during class hours). Two of the interventions (Bond et al., 2004; Holen, Waaktaar, Lervåg, & Ystgaard, 2012) involved moderate training in length, requiring attendance at a workshop of 2 days or less. Finally, one intervention (the Positive Adolescent Training through Holistic Social Programmes [PATHS] project; Shek & Ma, 2012) involved little to no training reflected in higher ratings of transferability.

The ultimate goal of EI interventions is to improve students’ self and social awareness, empathy, and perspective-taking ability, all factors that contribute to positive mental health. The most effective interventions involved the parents/family at some level to help maintain the students’ progress at home as well as school. Parental involvement ranged from minimal (e.g., receiving handouts detailing what their child has learned) to more extensive (e.g., parental participation in sessions outside of the classroom). Effective interventions in this category were primarily psychoeducational in nature, designed to promote emotional literacy to effectively combat a wide range of problems (e.g., bullying, substance abuse, behavioural problems, etc.). For example, the RULER Approach (Bond et al., 2004) involves Recognizing, Understanding, Labeling, Expressing, and Regulating emotions to promote effective social and emotional learning within the classroom.

Table 1 provides a summary of studies related to emotional intelligence.

SR

The reduction of stigma in mental health has as its focus reducing the shame associated with a mental health disorder. It has become common for media campaigns to
Table 1. Summary of Studies Related to Emotional Intelligence.

<table>
<thead>
<tr>
<th>Reference</th>
<th>Intervention category</th>
<th>Specific intervention</th>
<th>Age range</th>
<th>Effect size</th>
<th>Quality rating</th>
<th>Strength rating</th>
<th>Overall rating</th>
</tr>
</thead>
</table>
| S. M. Jones, Brown, and Lawrence Aber (2011)        | Modified RCT          | 4Rs Program           | 8-10 years          | ES\textsubscript{range} = 0.13-0.42  
ES\textsubscript{mean} = 0.19 | 3              | 1              | 3.1            |
| Shek and Ma (2012)                                  | Modified RCT          | PATHS Project         | 6-10 years          | ES\textsubscript{range} = 0.00-0.36  
ES\textsubscript{mean} = 0.19 | 3              | 1              | 3.1            |
| Li et al. (2011)                                    | Modified RCT          | Positive Action       | 8-11 years          | ES\textsubscript{range} = 0.27-0.41  
ES\textsubscript{mean} = 0.34 | 3              | 2              | 3.2            |
| Griggs, Rimm-Kaufman, Merritt, and Patton (2013)    | RCT                   | Responsive Classroom  | 10-11 years         | Not reported                | 4              | 1              | 4.1            |
| Bond et al. (2004)                                  | Modified RCT          | RULER Approach        | 10-12 years         | ES\textsubscript{range} = 0.50-0.60  
ES\textsubscript{mean} = 0.54 | 3              | 2              | 3.2            |
| Kimber, Sandell, and Bremberg (2008)                | Quasi                 | SET Program           | 9-15 years          | ES\textsubscript{range} = 0.11-0.44  
ES\textsubscript{mean} = 0.31 | 2              | 2              | 2.1            |
| Brown, Low, Smith, and Haggerty (2011)              | Modified RCT          | Steps to Respect      | 8-11 years          | ES\textsubscript{range} = 0.12-1.27  
ES\textsubscript{mean} = 0.36 | 3              | 2              | 3.2            |
| Bond et al. (2004)                                  | RCT                   | The Gatehouse Project | 13-16 years         | Not reported                | 3              | 1              | 3.1            |
| Fitzpatrick et al. (2013)                           | RCT                   | Working Things Out    | 12-17 years         | ES\textsubscript{range} = 0.13-0.28  
ES\textsubscript{mean} = 0.17 | 4              | 1              | 4.1            |
| Holen, Waaktaar, Lervåg, and Ystgaard (2012)        | Modified RCT          | Zippy’s Friends       | 7-8 years           | d\textsubscript{range} = 0.15-0.55  
\textsubscript{mean} = 0.31 | 3              | 2              | 3.2            |

Note: Quality scores range from 1 (weak) to 4 (strong), and strength 1 = weak or 2 = strong. RCT = randomized clinical trial; RULER = Recognizing, Understanding, Labeling, Expressing, and Regulating; SET = Social Emotional Learning Program.
feature well-known personalities who discuss their own struggles with a disorder. The goal of these campaigns is based on a premise that individuals will be more inclined to seek help if they do not suffer the secondary effects of negative attitudes and discrimination associated with what it means to have a mental health disorder. The Canadian Medical Association (2008) reported that only half of Canadians would feel comfortable discussing the fact of a personal mental issue with their family, in contrast to the 68% who would freely disclose a diagnosis of cancer (Canadian Medical Association, 2008).

Attitudes related to such mental health stigma appear to be well established at even a young age (Heflinger & Hinshaw, 2010). The Ontario Centre of Excellence in Child and Youth Mental Health (2014) suggested that in counteracting the effects of stigma, “There is growing evidence that schools, with their central role in the lives of young people, families and communities, are an ideal setting to start shaping a world without stigma” (p. 7).

The quality of the evidence for SR was comprised of six RCTs and four quasi-experimental studies, with ratings ranging from low to high. The sample sizes ranged from 92 to 783 participants. The mean age of participants for all 10 SR programs was 13.7 years ($SD = 1.74$). The average program duration was 7.1 days ($SD = 17.28$).

Although the majority of programs were manualized, four interventions (Conrad et al., 2009; Economou et al., 2012; Rickwood, Cavanagh, Curtis, & Sakrouge, 2004; Shryane, Byrne, Morrison, & Campbell, 2011) were semi-structured, involving speeches by and discussion (i.e., Q&A) with survivors of mental illness. In these programs, interventions were given a weaker recommendation, as they could not be implemented easily by all members of the school staff due to some not having access to a mental health survivor. The goal of involving a survivor of mental illness was meant to reduce stigma drawing on the contact hypothesis (Allport, 1954) which suggests that contact with the survivor could reduce prejudice or stigma toward the associated mental illness by facilitating a positive interaction and providing opportunities for empathy with these individuals. “One off” programs need to be used with caution to ensure that the speech is drawing from evidence-based information and that a clinician is present to address issues that may be raised by vulnerable students (Mental Health Commission of Canada, 2013b). These types of programs are very popular in schools with limited scrutiny, if any, being paid to the potential concerns.

All of the SR interventions were short in duration, most (70%) were only a day in length and therefore intervention frequency was not applicable. However, the frequencies of those longer than a day ranged from 1 (Ranson & Byrne, 2014) to 3 times a week (Wahl, Susin, Kaplan, Lax, & Zatina, 2011). Sixty percent of the interventions were manualized, while 40% were semistructured. Unfortunately, half of the studies did not report information on program training, as researchers implemented the program themselves. One program (Conrad et al., 2009) involved moderate training requiring attendance at a workshop of 2 days or less, while the remaining 40% of the interventions involved little to no training allowing for relatively high levels of transferability.

Effective interventions in this category were primarily psychoeducational in nature, with interactive components through group discussion and question and answer. One
intervention (Saporito, Ryan, & Teachman, 2011) combined an interactive PowerPoint presentation with a brief video of an adolescent struggling with mental illness. This intervention was particularly effective, as it not only used the contact hypothesis to reduce stigma, but the video format of this “contact” allowed for a higher degree of transferability than a semi-structured “live” presentation. Finally, similar to other effective intervention categories, the most effective interventions involved a homework component. For example, following each session of Understanding Our Peers (Ranson & Byrne, 2014), participants were asked to complete an online activity on a web-based learning site as homework. This was done to ensure that students thought about the material learned both inside and outside of the classroom.

**Mindfulness School-Based Program (MSBP)**

Mindfulness as a construct of thinking reflects both styles and content and has become extremely popular as reflected in both the scientific and mainstream literatures. Mindfulness is generally concerned with self and therapeutic training that is focused on attending to the present, being aware of the immediate context and letting go of personal judgments regarding categorical objections to others, that is, good or bad. Greenberg and Harris (2012) indicated there are literally thousands of sites extolling the virtues of mindfulness for children and youth, many of which appear in the context of preventive efforts to teach children to learn the strategies to counteract the impact of physical and mental health challenges.

Reviews on MSBP reflected 10 studies, half of which utilized a randomized control design with the balance utilizing a quasi-experimental design. The sample sizes ranged from 18 to 393 participants. The mean age of participants for all 10 mindfulness programs was 12.7 years (SD = 3.31). Average program duration was 9.9 weeks (SD = 4.68). Programs having a longer duration were downgraded due to possible difficulties in transferability to the school environment which included the Integra Mindfulness Martial Arts Program (Haydicky, Wiener, Badali, Milligan, & Ducharme, 2012). The majority of these 10 programs, including the BREATHE (Broderick & Metz, 2009), a mindfulness-based training program designed to facilitate the development of emotion regulation and attentional skills for middle and high school students and the Mindfulness Awareness Programs (MAPs) (Flook et al., 2010) programs, had a goal to teach students the skill of mindfulness. Programs, such as the Yoga Education program (Steiner, Sidhu, Pop, Frenette, & Perrin, 2013), teach students specific mindfulness skills to reduce emotional and/or behavioural disorders. The Integra Mindfulness Martial Arts Program was specifically designed for students who have learning disabilities and co-occurring ADHD, anxiety, and depression. The majority (70%) of these mindfulness programs were manualized and hence received higher overall ratings due to their ease of transferability to a school environment. The majority of these programs were preventive in teaching students mindfulness skills prior to the development of negative emotional or behavioural tendencies. Table 2 provides a summary of studies related to stigma reduction.

Overall, these studies had a mean effect size of .34 (SD = .19). Following from Cohen’s (1988) effect size guidelines, this indicates the MSBP programs reflect an
<table>
<thead>
<tr>
<th>Reference</th>
<th>Intervention category</th>
<th>Specific intervention</th>
<th>Age range</th>
<th>Effect size</th>
<th>Quality rating</th>
<th>Strength rating</th>
<th>Overall rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conrad et al. (2009)</td>
<td>Quasi</td>
<td>“Crazy? So what!”</td>
<td>13-18 years 8th-12th grade</td>
<td>Not reported</td>
<td>1</td>
<td>1</td>
<td>1.1</td>
</tr>
<tr>
<td>Chan, Mak, and Law (2009)</td>
<td>RCT</td>
<td>“The Same or Not the Same”</td>
<td>14-15 years 9th grade</td>
<td><strong>ES</strong> range = 0.24-0.50</td>
<td>3</td>
<td>2</td>
<td>3.2</td>
</tr>
<tr>
<td>Wahl, Susin, Kaplan, Lax, and Zatina (2011)</td>
<td>Quasi</td>
<td>Breaking the Silence</td>
<td>12-14 years 7th-8th grade</td>
<td>Not reported</td>
<td>2</td>
<td>1</td>
<td>2.1</td>
</tr>
<tr>
<td>Economou et al. (2012)</td>
<td>RCT</td>
<td>Unnamed Stigma Reduction Program</td>
<td>13-15 years 8th-9th grade</td>
<td>Not reported</td>
<td>4</td>
<td>1</td>
<td>4.1</td>
</tr>
<tr>
<td>Rickwood, Cavanagh, Curtis, and Sakrouge (2004)</td>
<td>Quasi</td>
<td>Mental Illness Education Program</td>
<td>14-18 years 9th-12th grade</td>
<td><strong>ES</strong> mean = 0.25</td>
<td>1</td>
<td>1</td>
<td>1.1</td>
</tr>
<tr>
<td>Shryane, Byrne, Morrison, and Campbell (2011)</td>
<td>RCT</td>
<td>Psychosocial Mental Health Promotion</td>
<td>14-15 years 9th grade</td>
<td>ES = 0.30</td>
<td>3</td>
<td>2</td>
<td>3.2</td>
</tr>
<tr>
<td>Martiniuk, Speechley, Secco, Campbell, and Donner (2007)</td>
<td>RCT</td>
<td>Thinking About Epilepsy</td>
<td>10-11 years 5th grade</td>
<td><strong>ES</strong> range = 0.28-0.63, <strong>ES</strong> mean = 0.46</td>
<td>4</td>
<td>2</td>
<td>4.2</td>
</tr>
<tr>
<td>Ranson and Byrne (2014)</td>
<td>Modified RCT</td>
<td>Understanding Our Peers</td>
<td>12-13 years 8th grade</td>
<td>(d) range = 0.32-1.37, (d) mean = 0.72, (r) range = 0.16-0.56, (r) mean = 0.32</td>
<td>3</td>
<td>2</td>
<td>3.2</td>
</tr>
<tr>
<td>Ventieri, Clarke, and Hay (2011)</td>
<td>Quasi</td>
<td>Unnamed Stigma Reduction Program</td>
<td>10-12 years 5th-6th grade</td>
<td>(d) range = 0.25-0.73, (d) mean = 0.44, (r) range = 0.12-0.34, (r) mean = 0.21</td>
<td>2</td>
<td>2</td>
<td>2.2</td>
</tr>
<tr>
<td>Saporito, Ryan, and Teachman (2011)</td>
<td>RCT</td>
<td>Unnamed Stigma Reduction Program</td>
<td>15-19 years 10th-12th grade</td>
<td>(d) range = 0.64-0.97, (d) mean = 0.80</td>
<td>3</td>
<td>2</td>
<td>3.2</td>
</tr>
</tbody>
</table>

Note. Quality scores range from 1 (weak) to 4 (strong) and strength 1 = weak or 2 = strong. RCT = randomized clinical trial.
overall small effect size. In accordance with the GRADE analysis, the quality of the research literature ranges from very low to high. Most (40%) of these studies had a high quality and 30% had a low quality. For strength of research, 20% of the studies were considered strong as reflected in their strength of research and 80% were weak. The overall rating of studies ranged from very low-weak to high-strong. Of the 10 studies reviewed, one program, MindUp (Schonert-Reichel & Lawlor, 2010), had a high-strong overall rating. Three studies (30%) received an overall rating of high-weak.

Mindfulness programs had an average small effect size. Four of the studies had a high quality with the majority having a weak strength of research. The 10 different interventions taught children a variety of skills such as body awareness (the BREATHE program), awareness of others (the MAPs program), and emotional awareness (the Mindfulness Meditation program). However, the more effective programs that received a high-strong or high-weak overall rating encouraged awareness of the present time such as the Mindfulness-Based Stress Reduction Program and perspective taking skills as displayed in the MindUp program.

**Anxiety Programs**

Costello, Egger, and Angold (2005) reported that anxiety disorders are among the most common mental health challenges facing young people. Anxiety in childhood typically refers to an over concern related to personal well-being. This can be reflected in excessive worrying, nervousness, or fear. Such behavioural or emotional expressions are experienced by most children and youth as part of a developmental progression prompted by changes in stage development leading to greater personal security. There is concern for certain anxiety disordered youth, however, reflected in a stronger than usual or over generalization of anxiety responses that becomes debilitating, interfering with normal functioning such as attendance at school, fear of separation from a primary caregiver, or other such functional impairments (Kendall, Compton, et al., 2010).

The anxiety research literature summarizes 10 programs. Eight of these 10 studies were RCTs with the remaining two utilizing quasi-experimental designs. The sample sizes ranged from 40 to 1,477 participants. The mean age of participants for all 10 anxiety prevention programs was 12.2 years \((SD = 2.31)\). The average program length was 16.1 weeks \((SD = 11.16)\). Programs that were lengthier were downgraded due to possible difficulties in transferability to a school environment such as the Psycho-Educational Program (Balle & Tortella-Feliu, 2010). Eight of the 10 programs targeted general anxiety. One article specifically targeted test anxiety, the “Clicker” Response System program (Agarwal, D’Antonio, Roediger, McDermott, & McDaniel, 2014), with the Confident Kids program (Berry & Hunt, 2009) targeting anxiety resulting from bullying. Half of the anxiety prevention programs were manualized and required no formalized training on behalf of the administrators. The other half of the programs required attendance at a workshop lasting no longer than 2 days. The programs that required more administrator training were downgraded. Half of the programs focused on prevention in implementing an intervention prior to the development of anxiety symptoms.
Table 3 provides a summary of studies related to the implementation of mindfulness programs. These studies had a mean effect size of .40 ($SD = .29$). Following from Cohen’s (1988) effect size guidelines, this indicates that the anxiety school-based prevention program literature possess a small effect size. In accordance with the GRADE analysis, the quality of the research literature ranges from very low to high. Half of these studies have a high quality rating. For strength of research, 30% of the programs were rated high in their strength of research. The overall rating of studies ranged from very low-weak to high-weak. Of the 10 programs reviewed, two had a high-strong overall rating. Half of the studies received an overall rating of high-weak. The downgrading of studies was due to a lack of generalizability of the program such as with the MoodGym program, a lack of data analysis (the “Clicker” Response System program), and lack of significance for certain outcome measures such as the Physical Activity Program (Bonhauser et al., 2005).

The 10 different interventions addressed a variety of anxiety symptoms such as low self-esteem in the Confident Kids program, maladaptive coping skills in the Feelings Club Program (Manassis et al., 2010), and emotional distress, with the Physical Activity Program. However, the highly effective programs encouraged the development of positive thinking as with the FRIENDS program, and in the development of coping skills such as with the Skills for Positive Thinking Program (Cunningham, Brandon, & Frydenberg, 2002).

**Depression**

Although depression in childhood and adolescence can be similar in presentation to that identified in adults, Bhatia and Bhatia (2007) suggested that many children and young adults will not manifest the emotional content of depression but will give evidence of their distress through “vague somatic symptoms or pain, irritable mood, anger, hostile behaviour, impaired attention, poor concentration . . . attention-deficit/ hyperactivity disorder, and substance abuse” (p. 75). It is estimated that while at any given time 1% to 4% of youth experience a major depressive disorder, many more youth experience transient feelings of sadness and dysphoria and that the onset of depression in adolescence often marks the beginning of increases in the prevalence rate throughout early adulthood (Merikangas, Nakamura, & Kessler, 2009).

The depression program evidence was comprised of 10 RCTs, ranging from moderate to high strength, with sample sizes from 120 to 1,477 participants. The mean age of participants was 14.1 years ($SD = 2.19$). The average program duration was 9.5 weeks ($SD = 2.92$). Six out of 10 studies demonstrated a moderate-to-strong positive effect. However, in one study (Pössel, Seemann, & Hautzinger, 2008), depressive symptoms actually increased significantly more in the intervention group than the control group for boys suggesting a potentially harmful effect from the intervention. As a result, this particular intervention was given a weak recommendation. Similarly, although the MoodGYM intervention (Calear, Christensen, Mackinnon, Griffiths, & O’Kearney, 2009) targeted both anxiety and depression, the intervention was found to be less effective in alleviating depression than anxiety, and was therefore given a weak
Table 3. Summary of Studies Related to Mindfulness.

<table>
<thead>
<tr>
<th>Reference</th>
<th>Intervention category</th>
<th>Specific intervention</th>
<th>Program appropriate for grades</th>
<th>Effect size</th>
<th>Quality rating</th>
<th>Strength rating</th>
<th>Overall rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Broderick and Metz (2009)</td>
<td>Quasi-experimental</td>
<td>BREATHE</td>
<td>Study and Program Grades: 12</td>
<td>ES&lt;sub&gt;range&lt;/sub&gt; = 0.06-0.57</td>
<td>2</td>
<td>1</td>
<td>2.1</td>
</tr>
<tr>
<td>Flook et al. (2010)</td>
<td>RCT</td>
<td>MAPs</td>
<td>Study and Program Grades: 2-3</td>
<td>d&lt;sub&gt;mean&lt;/sub&gt; = 0.28</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Haydicky, Wiener, Badali, Milligan, and Ducharme (2012)</td>
<td>Quasi-experimental</td>
<td>Integra Mindfulness Martial Arts Study and Program</td>
<td>Graded: 7-12</td>
<td>ES&lt;sub&gt;range&lt;/sub&gt; = 0.00-0.39</td>
<td>2</td>
<td>1</td>
<td>2.1</td>
</tr>
<tr>
<td>Liehr and Diaz (2010)</td>
<td>RCT</td>
<td>Mindful Schools</td>
<td>Study and Program Grades: 4</td>
<td>d&lt;sub&gt;mean&lt;/sub&gt; = 0.39</td>
<td>3</td>
<td>1</td>
<td>3.1</td>
</tr>
<tr>
<td>Raes, Griffith, Van der Gucht, and Williams (2014)</td>
<td>RCT</td>
<td>Mindfulness Intervention Program</td>
<td>Study and Program Grades: 8-12</td>
<td>ES&lt;sub&gt;range&lt;/sub&gt; = 0.05-0.32</td>
<td>4</td>
<td>1</td>
<td>4.1</td>
</tr>
<tr>
<td>Schonert-Reichl and Lawlor (2010)</td>
<td>Quasi-experimental</td>
<td>Mindfulness Education Program</td>
<td>Study and Program Grades: 4-7</td>
<td>ES&lt;sub&gt;range&lt;/sub&gt; = 0.45-0.97</td>
<td>2</td>
<td>2</td>
<td>2.2</td>
</tr>
<tr>
<td>Schonert-Reichl et al. (2015)</td>
<td>RCT</td>
<td>MindUp</td>
<td>Study and Program Grades: 4-5</td>
<td>ES&lt;sub&gt;range&lt;/sub&gt; = 0.17-0.87</td>
<td>4</td>
<td>2</td>
<td>4.2</td>
</tr>
<tr>
<td>Sibinga et al. (2013)</td>
<td>RCT</td>
<td>Mindfulness-Based Stress Reduction</td>
<td>Study and Program Grades: 7-8</td>
<td>d&lt;sub&gt;mean&lt;/sub&gt; = 0.51</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Steiner, Sidhu, Pop, Frenette, and Perrin (2013)</td>
<td>Quasi-experimental</td>
<td>Yoga Education</td>
<td>Study and Program Grades: 4-5</td>
<td>ES&lt;sub&gt;range&lt;/sub&gt; = 0.14-1.15</td>
<td>1</td>
<td>1</td>
<td>1.1</td>
</tr>
<tr>
<td>Wisner (2014)</td>
<td>Quasi-experimental</td>
<td>Mindfulness Meditation</td>
<td>Study and Program Grades: 10-12</td>
<td>d&lt;sub&gt;mean&lt;/sub&gt; = 0.33</td>
<td>1</td>
<td>1</td>
<td>1.1</td>
</tr>
</tbody>
</table>

Note. Quality scores range from 1 (weak) to 4 (strong) and strength 1 = weak or 2 = strong. RCT = randomized clinical trial.
recommendation within the depression intervention category. Table 4 provides a summary of studies related to anxiety reduction.

Interventions for depression ranged from shorter term (4 weeks) to longer term (12 weeks), and were typically implemented once a week for 90-min sessions outside of class time. The *Penn Resiliency Program* (Gillham et al., 2007) as an example was implemented more intensively over 12 weeks, sessions being held twice per week for the first 3 weeks, and once per week thereafter. All interventions were manualized involving scripts and lesson plans to increase transferability or ease of implementation. Most of these programs (70%) involved more intensive training requiring attendance at a workshop of more than 2 days with some form of supervision to follow. Although this negatively affected their transferability, the need for increased training and supervision appeared necessary considering the severe and pervasive nature of depression and the sensitivity and level of expertise needed to effectively address the topic. In contrast, two of the interventions (Kowalenko et al., 2005; McCarty, Violette, Duong, Cruz, & McCauley, 2013) involved moderate training with attendance at a workshop of 2 days or less. Finally, the *MoodGYM* program (Calear et al., 2009) involved little to no training and could be easily implemented within any classroom.

The most effective interventions were cognitive behavioural and psychoeducational in nature. For example, the *Coping With Stress Course* (Horowitz, Garber, Ciesla, Young, & Mufson, 2007) taught students how to (a) monitor daily moods; (b) identify activating events; (c) recognize, evaluate, challenge, and modify negative beliefs; (d) notice connections among activating events, beliefs, and consequences; and (e) cope with stressful events (p. 697). These interventions were typically conducted with smaller groups of students allowing for regular group discussion. For example, the *Penn Resiliency Program* (Gillham et al., 2007) used groups of 6-14 students which were designed to promote group cohesion, social support, discussion of depression relevant topics, sharing of thoughts/feelings, and engagement in activities through role-playing. Finally, typical of cognitive-behavioural interventions, the most effective interventions used a homework component to ensure that what the students learned was being practiced and applied outside of the classroom. Table 5 provides a summary of studies related to depression management.

**Addictions**

The Centre for Addiction and Mental Health’s (CAMH, 2015) most recent Ontario Student Drug Survey indicated that while certain drugs are showing a decrease in their use, such as tobacco, cannabis, and opioids, some drug use is still characteristic of well over 50% of students in Grades 7 through 12 as reflected in the four leading areas in the use of alcohol (45.8%), energy—high caffeinated drinks (34.8%), cannabis (21.3%), and binge drinking (17.6%).

The addictions research summarizes 10 programs, all of which were preventive in nature. Eight of these 10 studies utilized a RCT with the remaining two studies using a quasi-experimental design. The sample sizes ranged from 478 to 9,528 participants. The mean age of participants for all 10 programs was 12.2 years ($SD = 1.68$). The
<table>
<thead>
<tr>
<th>Reference</th>
<th>Intervention category</th>
<th>Specific intervention</th>
<th>Program appropriate for grades</th>
<th>Effect size</th>
<th>Quality rating</th>
<th>Strength rating</th>
<th>Overall rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agarwal, D’Antonio, Roediger, McDermott, and McDaniell (2014)</td>
<td>Quasi-experimental</td>
<td>“Clicker” Response System</td>
<td>Study and Program Grades: 6-12</td>
<td>Not reported</td>
<td>1</td>
<td>1</td>
<td>1.1</td>
</tr>
<tr>
<td>Balle and Tortella-Feliu (2010)</td>
<td>RCT</td>
<td>Psycho-Educational Program</td>
<td>Study and Program Grades: 6-12</td>
<td>$ES_{range} = 0.01-0.57$ $d_{mean} = 0.20$</td>
<td>4</td>
<td>1</td>
<td>4.1</td>
</tr>
<tr>
<td>Berry and Hunt (2009)</td>
<td>RCT</td>
<td>Confident Kids</td>
<td>Study and Program Grades: 7-10</td>
<td>$ES_{range} = -0.4-2.07$ $d_{mean} = 0.92$</td>
<td>3</td>
<td>2</td>
<td>3.2</td>
</tr>
<tr>
<td>Bonhauser et al. (2005)</td>
<td>Quasi-experimental</td>
<td>Physical Activity Program</td>
<td>Study and Program Grades: 9</td>
<td>Not reported</td>
<td>2</td>
<td>1</td>
<td>2.1</td>
</tr>
<tr>
<td>Calear, Christensen, Mackinnon, Griffiths, and O’Kearney (2009)</td>
<td>RCT</td>
<td>MoodGYM</td>
<td>Study and Program Grades: 7-12</td>
<td>$ES_{range} = 0.02-0.27$ $d_{mean} = 0.18$</td>
<td>4</td>
<td>1</td>
<td>4.1</td>
</tr>
<tr>
<td>Chiu et al. (2013)</td>
<td>RCT</td>
<td>Building Confidence</td>
<td>Study and Program Grades: 1-7</td>
<td>$ES_{range} = 0.28-1.62$ $d_{mean} = 0.74$</td>
<td>3</td>
<td>2</td>
<td>3.2</td>
</tr>
<tr>
<td>Cunningham, Brandon, and Frydenberg (2002)</td>
<td>RCT</td>
<td>Skills for Positive Thinking</td>
<td>Study and Program Grades: 5-6</td>
<td>$ES_{range} = 0.00-0.57$ $d_{mean} = 0.26$</td>
<td>4</td>
<td>1</td>
<td>4.1</td>
</tr>
<tr>
<td>Essau, Conradt, Sasagawa, and Ollendick (2012)</td>
<td>RCT</td>
<td>FRIENDS</td>
<td>Study and Program Grades: 4-12</td>
<td>$ES_{range} = 0.02-0.25$ $d_{mean} = 0.21$</td>
<td>4</td>
<td>1</td>
<td>4.1</td>
</tr>
<tr>
<td>Manassis et al. (2010)</td>
<td>RCT</td>
<td>The Feelings Club</td>
<td>Study and Program Grades: 3-6</td>
<td>$ES_{range} = 0.06-0.44$ $d_{mean} = 0.17$</td>
<td>4</td>
<td>1</td>
<td>4.1</td>
</tr>
<tr>
<td>March, Spence, and Donovan (2009)</td>
<td>RCT</td>
<td>BRAVE</td>
<td>Study and Program Grades: 2-7</td>
<td>$ES_{range} = 0.31-0.77$ $d_{mean} = 0.53$</td>
<td>3</td>
<td>2</td>
<td>3.2</td>
</tr>
</tbody>
</table>

*Note. Quality scores range from 1 (weak) to 4 (strong) and strength 1 = weak or 2 = strong. RCT = randomized clinical trial.*
<table>
<thead>
<tr>
<th>Reference</th>
<th>Intervention category</th>
<th>Specific intervention</th>
<th>Age range</th>
<th>Effect size</th>
<th>Quality</th>
<th>Strength</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kowalenko et al. (2005)</td>
<td>Modified RCT</td>
<td>ACE Program</td>
<td>13-16 years</td>
<td>$d_{\text{range}} = 0.12-0.57$ $d_{\text{mean}} = 0.43$ $r_{\text{range}} = 0.06-0.27$ $r_{\text{mean}} = 0.21$</td>
<td>3</td>
<td>2</td>
<td>3.2</td>
</tr>
<tr>
<td>Stice, Burton, Bearman, and Rohde (2007)</td>
<td>RCT</td>
<td>Blues Group</td>
<td>15-22 years</td>
<td>$d_{\text{range}} = 0.14-0.71$ $d_{\text{mean}} = 0.42$ $r_{\text{range}} = 0.7-0.33$ $r_{\text{mean}} = 0.20$</td>
<td>3</td>
<td>2</td>
<td>3.2</td>
</tr>
<tr>
<td>Horowitz, Garber, Ciesla, Young, and Mufson (2007)</td>
<td>RCT</td>
<td>Coping with Stress Course</td>
<td>14-15 years</td>
<td>$d_{\text{range}} = 0.37-0.89$ $d_{\text{mean}} = 0.63$</td>
<td>4</td>
<td>2</td>
<td>4.2</td>
</tr>
<tr>
<td>Pössel et al., (2008)</td>
<td>RCT</td>
<td>LARS &amp; LISA(^a)</td>
<td>13-14 years</td>
<td>$g_{\text{range}} = 0.43-0.61$ $g_{\text{mean}} = 0.55$</td>
<td>4</td>
<td>1</td>
<td>4.1</td>
</tr>
<tr>
<td>Calear et al. (2009)</td>
<td>RCT</td>
<td>MoodGYM</td>
<td>14-18 years</td>
<td>$d_{\text{range}} = 0.27-0.43$</td>
<td>4</td>
<td>1</td>
<td>4.1</td>
</tr>
<tr>
<td>Gillham et al. (2007)</td>
<td>RCT</td>
<td>Penn Resiliency Program</td>
<td>13-15 years</td>
<td>$ES_{\text{range}} = 0.18-0.58$ $ES_{\text{mean}} = 0.38$</td>
<td>4</td>
<td>2</td>
<td>4.2</td>
</tr>
<tr>
<td>McCarty, Violette, Duong, Cruz, and McCauley (2013)</td>
<td>RCT</td>
<td>PTA Program</td>
<td>13-15 years</td>
<td>$d_{\text{range}} = 0.36-0.54$ $d_{\text{mean}} = 0.43$</td>
<td>3</td>
<td>2</td>
<td>3.2</td>
</tr>
<tr>
<td>Rivet-Duval, Heriot, and Hunt (2011)</td>
<td>RCT</td>
<td>Resourceful Adolescent Program</td>
<td>12-15 years</td>
<td>$d_{\text{range}} = 0.32-0.67$ $d_{\text{mean}} = 0.43$</td>
<td>3</td>
<td>2</td>
<td>3.2</td>
</tr>
<tr>
<td>Cabiya et al. (2008)</td>
<td>RCT</td>
<td>Unnamed CBT Program</td>
<td>13-15 years</td>
<td>$ES_{\text{range}} = 0.06-1.02$ $ES_{\text{mean}} = 0.46$</td>
<td>4</td>
<td>2</td>
<td>4.2</td>
</tr>
<tr>
<td>Arnarson and Craighead (2009)</td>
<td>RCT</td>
<td>Unnamed Program</td>
<td>14-15 years</td>
<td>$OR = 0.12$</td>
<td>3</td>
<td>1</td>
<td>3.1</td>
</tr>
</tbody>
</table>

**Table 5. Summary of Studies Related to Depression.**

Note. Quality scores range from 1 (weak) to 4 (strong) and strength 1 = weak or 2 = strong. RCT = randomized clinical trial; ACE = Adolescents Coping with Emotions; PTA = Positive Thoughts and Actions; CBT = Cognitive Behavioural Program.

\(^a\)LARS & LISA are gender aware programs that promote understanding of the key concepts through exploring the relationship between cognition, emotion, and behavior.
average program duration was 23.3 weeks ($SD = 15.11$). Seventy percent of these programs targeted general substance abuse while 30% specifically targeted alcohol abuse. The alcohol abuse programs included the *CHOICE* program (D’Amico et al., 2012), the *Parent and Student Intervention* program (Koning, Van den Eijnden, Verdurmen, Engels, & Vollebergh, 2011), and the *Climate Schools* program (Newton, Teesson, Vogl, & Andrews, 2010). The *Village Model of Care* program (Hanlon, Simon, O’Grady, Carswell, & Callaman, 2009) spoke to the development of an addiction prevention program specifically for youths of African American decent. With the exception of one program, all programs were manualized and required no formal training.

The programs yielded a mean effect size of $.65$ ($SD = .26$) reflecting an overall medium effect size (Cohen, 1988). Program quality ranged from low to high, with 70% of the studies reflecting a high quality. For strength of research, 60% of the programs were rated strong in their strength of research with the balance having a low strength. The overall rating of studies ranged from low-weak to high-strong. Half of these programs had a high-strong overall rating. The downgrading of studies was generally due to a high attrition rate such as in *Life Skills Program* (Wenzel, Weichold, & Silbereisen, 2009) a lack of significance for certain outcome measures reflected in the *Village Model of Care* and the possibility of a response bias reflected in the *Parent and Student Intervention Program*.

The 10 different interventions targeted the development of a variety of skills in students such as maintaining a positive self-concept in *Unplugged* (Faggiano et al., 2010), self-efficacy in *Media Detective Skills* (Kupersmidt, Scull, & Austin, 2010), and problem-solving skills in the *Project Toward No Drug Abuse* program (Rohrbach, Gunning, Sun, & Sussman, 2010) and communication skills reflected in the *European Drug Abuse Program* (Vigna-Taglianti et al., 2009). The most effective programs taught students about the negative effects of substance abuse such as in the *Media Detective Skills* program and *Project Toward No Drug Abuse* program. In general, these addiction prevention programs are very effective, highly accessible (i.e., “off the shelf”), and are readily transferable to a school environment. Table 6 provides a summary of studies related to addiction.

**Suicide Prevention**

Suicide in adolescence is now considered a public health concern as it represents the third leading cause of death among young people (Cash & Bridge, 2009). Trends in Canada reflect increasing rates of suicide among female children and adolescents and decreases among equivalent males (Skinner & McFaul, 2012). Many of the girls who reported a suicide attempt had presented to a hospital with a mental illness in the preceding year (Rhodes, Bethell, et al., 2014). The American Psychological Association (2015) indicated that suicide prevention efforts have

focused on school education programs, crisis center hotlines, screening programs that seek to identify at-risk adolescents, media guidelines (suicide prevention strategies that
<table>
<thead>
<tr>
<th>Reference</th>
<th>Intervention category</th>
<th>Specific intervention</th>
<th>Program appropriate for grades</th>
<th>Effect size</th>
<th>Quality rating</th>
<th>Strength rating</th>
<th>Overall rating</th>
</tr>
</thead>
</table>
| Beets et al. (2009) | RCT | Positive action | Study Grades: 1-5 | $ES_{\text{range}} = 0.12-1.49$  
$ES_{\text{mean}} = 0.82$ | 4 | 2 | 4.2 |
| D’Amico et al. (2012) | RCT | CHOICE | Study and Program Grades: 6-8 | $ES_{\text{range}} = 0.7-1.11$  
$ES_{\text{mean}} = 0.81$ | 3 | 2 | 3.2 |
| Faggiano et al. (2010) | RCT | Unplugged | Study and Program Grades: 7-9 | $ES_{\text{range}} = 0.62-0.94$  
$ES_{\text{mean}} = 0.82$ | 4 | 2 | 4.2 |
| Hanlon, Simon, O’Grady, Carswell, and Callaman (2009) | Quasi-Experimental | The Village Model of Care | Study and Program Grades: 6 | $ES_{\text{range}} = 0.04-0.38$  
$ES_{\text{mean}} = 0.22$ | 2 | 1 | 2.1 |
| Koning, Van den Eijnden, Verdurmen, Engels, and Vollebergh (2011) | RCT | Parent and student intervention | Study and Program Grades: 7 | $ES_{\text{range}} = 0.59-0.88$  
$ES_{\text{mean}} = 0.76$ | 3 | 1 | 3.1 |
| Kupersmidt, Scull, and Austin (2010) | RCT | Media detective | Study and Program Grades: 2-8 | $ES_{\text{range}} = 0.96-0.98$  
$ES_{\text{mean}} = 0.97$ | 4 | 2 | 4.2 |
| Newton, Teesson, Vogl, and Andrews (2010) | RCT | Climate schools | Study and Program Grades: 8 | $ES_{\text{range}} = 0.1-0.81$  
$ES_{\text{mean}} = 0.36$ | 4 | 1 | 4.1 |
| Rohrbach, Gunning, Sun, and Sussman (2010) | RCT | Project toward no drug abuse | Study and Program Grades: 8-12 | $ES_{\text{range}} = 0.17-0.85$  
$ES_{\text{mean}} = 0.62$ | 4 | 2 | 4.2 |
| Vigna-Taglianti et al. (2009) | RCT | European drug abuse program | Study and Program Grades: 6-12 | $ES_{\text{range}} = 0.49-1.4$  
$ES_{\text{mean}} = 0.83$ | 4 | 2 | 4.2 |
| Wenzel, Weichold, and Silbereisen (2009) | Quasi-Experimental | Life Skills Program | Study and Program Grades: 5-8 | $ES_{\text{range}} = 0.21-0.41$  
$ES_{\text{mean}} = 0.31$ | 2 | 1 | 2.1 |

Note. RCT = randomized clinical trial.
involve educating media professionals about the prevalence of copy-cat suicides among adolescents, in an effort to minimize the impact of news stories reporting suicide) and efforts to limit firearm access. (p. 1)

There were eight school-based suicide prevention programs for children and youth that met the eligibility criteria for inclusion in this review. Of these eight programs, 75% utilized a RCT, one study used a quasi-experimental design and one study was observational. The sample sizes ranged from 128 to 3,128 participants. The mean age of participants was 14.75 years ($SD = 11.15$). The average program length was 18.6 weeks ($SD = 11.15$). The majority of these eight programs sought to teach students the ability to identify emotional distress as in the SOS program (Schilling, Lawless, Buchanan, & Aseltine, 2014), self-harm awareness with the Distress-Coping program (Klingman & Hochdorf, 1993), and to recognize the signs of potential suicide risk reflected in the South Elgin High School Suicide Prevention program (Ciffone, 2007). One program, the Zuni Life Skills program (LaFromboise & Howard-Pitney, 1995), was specifically designed for students of Native American decent. The majority of these studies (62.5%) required administrators to attend a workshop no longer than 2 days in length. Three programs were manualized and did not require training beyond reading the manual, and as a result received a higher GRADE rating due to ease of transportability to the school environment. The manualized programs included, the South Elgin High School Suicide Prevention program, the Distress-Coping program and the Zuni Life Skills program. All programs were designed to inform students regarding suicide awareness as well as those who may be at greater risk for suicide.

Overall, the programs had a mean effect size of .37 ($SD = .16$) reflecting an overall small effect size (Cohen, 1988). The quality of the studies ranged from very low to high, with 37.5% of these studies having a high quality and 37.5% having a moderate quality. The overall rating of articles ranged from very low-weak to high-weak. Of the eight articles reviewed, 37.5% had a high-strong overall rating. The programs receiving a high-weak rating were the Distress-Coping program, the Zuni Life Skills program, and the Sources of Strength program (Wyman et al., 2010). The more effective programs encouraged the development of protective factors against suicide such as informing someone of an “at risk” student and teaching students that suicide is never the best option. Table 7 provides a summary of studies related to suicide prevention.

**Trauma**

Concepts related to trauma reflect an individual’s inability to forget about a distressing event, psychological reliving of the event through triggers that remind the individual of past upsetting events, hyperarousal, and/or actual negative changes to thoughts and beliefs (Van der Kolk, 2014). Van der Kolk (2006) has advocated for adoption of the concept of developmental traumatology in furthering understanding regarding the unique impact of traumatic experience on children and adolescence who, by virtue of the context of their development at the time of the trauma, are challenged in processing the impact of a traumatic experience, thereby impacting their brain function and
<table>
<thead>
<tr>
<th>Reference</th>
<th>Intervention category</th>
<th>Specific intervention</th>
<th>Program appropriate for grades</th>
<th>Effect size</th>
<th>Quality rating</th>
<th>Strength rating</th>
<th>Overall rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ciffone (2007)</td>
<td>RCT</td>
<td>South Elgin High School Suicide Prevention Program</td>
<td>Study and Program Grades: 10</td>
<td>ES&lt;sub&gt;range&lt;/sub&gt; = 0.85-0.96 ES&lt;sub&gt;mean&lt;/sub&gt; = 0.9</td>
<td>3</td>
<td>2</td>
<td>3.2</td>
</tr>
<tr>
<td>Cigularov, Chen, Thurber, and Stallones (2008)</td>
<td>Quasi-Experimental</td>
<td>Raising Awareness of Personal Power</td>
<td>Study and Program Grades: 8-12</td>
<td>ES&lt;sub&gt;range&lt;/sub&gt; = 0.15-0.97 ES&lt;sub&gt;mean&lt;/sub&gt; = 0.5</td>
<td>1</td>
<td>2</td>
<td>1.2</td>
</tr>
<tr>
<td>King, Strunk, and Sorter (2011)</td>
<td>Observational</td>
<td>Surviving the Teens</td>
<td>Study and Program Grades: 9-12</td>
<td>ES&lt;sub&gt;range&lt;/sub&gt; = 0.07-0.36 ES&lt;sub&gt;mean&lt;/sub&gt; = 0.1</td>
<td>1</td>
<td>1</td>
<td>1.1</td>
</tr>
<tr>
<td>Klingman and Hochdorf (1993)</td>
<td>RCT</td>
<td>Distress-Coping</td>
<td>Study and Program Grades: 8</td>
<td>ES&lt;sub&gt;range&lt;/sub&gt; = 0.08 ES&lt;sub&gt;mean&lt;/sub&gt; = 0.01-0.17</td>
<td>4</td>
<td>1</td>
<td>4.1</td>
</tr>
<tr>
<td>LaFromboise and Howard-Pitney (1995)</td>
<td>RCT</td>
<td>Zuni Life Skills</td>
<td>Study and Program Grades: 9-12</td>
<td>ES&lt;sub&gt;range&lt;/sub&gt; = 0.2-0.51 ES&lt;sub&gt;mean&lt;/sub&gt; = 0.36</td>
<td>4</td>
<td>1</td>
<td>4.1</td>
</tr>
<tr>
<td>Portzky and Van Heeringen (2006)</td>
<td>RCT</td>
<td>Knowledge About Suicide</td>
<td>Study and Program Grades: 9-12</td>
<td>ES&lt;sub&gt;range&lt;/sub&gt; = 0.03-0.36 ES&lt;sub&gt;mean&lt;/sub&gt; = 0.2</td>
<td>3</td>
<td>1</td>
<td>3.1</td>
</tr>
<tr>
<td>Schilling, Lawless, Buchanan, and Aseltine (2014)</td>
<td>RCT</td>
<td>SOS</td>
<td>Study and Program Grades: 5-8</td>
<td>ES&lt;sub&gt;range&lt;/sub&gt; = 0.38-0.4 ES&lt;sub&gt;mean&lt;/sub&gt; = 0.39</td>
<td>3</td>
<td>1</td>
<td>3.1</td>
</tr>
<tr>
<td>Wyman et al. (2010)</td>
<td>RCT</td>
<td>Sources of Strength</td>
<td>Study and Program Grades: 10-11</td>
<td>ES&lt;sub&gt;range&lt;/sub&gt; = 0.11-0.63 ES&lt;sub&gt;mean&lt;/sub&gt; = 0.43</td>
<td>4</td>
<td>1</td>
<td>4.1</td>
</tr>
</tbody>
</table>

Note. Quality scores range from 1 (weak) to 4 (strong) and strength 1 = weak or 2 = strong. RCT = randomized clinical trial; SOS = Signs of Suicide.
interfering with their capacity to engage in the present. Bryant (2014) made the case that early intervention or intervention closer to the time of the traumatic event can help mediate the effects of the role played by shame, self-blame, emerging interpersonal difficulties, and the learning of avoidant coping strategies resulting from the trauma. Rolfsnes and Idose (2011) suggested that schools are a natural place to teach resiliency skills as they are familiar, safe, and supportive settings for children and young adults to learn resiliency strategies in the face of trauma.

The majority (63.63%) of the 11 trauma programs utilized randomized designs with the remainder being quasi-experimental in nature. Sample sizes ranged from 78 to 5,651 participants, with a mean age of 11.67 years ($SD = 2.3$). The average program duration was 20.7 weeks ($SD = 13.25$). Program components attended to an array of possible causes of trauma including political conflict as in the PSSA program (Ager et al., 2011), dating violence with Expect Respect program (Ball, Kerig, & Rosenbluth, 2009), and bullying in Steps to Respect program (Brown, Low, Smith, & Haggerty, 2011), The Modified FRIENDS program (Cooley-Strickland, Griffin, Darney, Otte, & Ko, 2011) was specifically designed for students of African American descent while the Mental Health Service program (Fazel, Doll, & Stein, 2009) was specifically designed for refugee children and youth. The majority (54.54%) of these programs were manualized and required no formal training beyond reading the manual for administration.

These programs yielded effect sizes ranging from low to highly effective with a mean effect size of .53 ($SD = .44$) suggesting a medium effect size (Cohen, 1988). The quality of the studies ranged from very low to high. One third (36.36%) had a high quality, with 36.4% having a moderate quality. Less than half (45.5%) of these programs were rated as strong with the balance having a rating of weak. The overall rating of articles ranged from very low/weak to high/weak. Slightly more than a quarter (27.3%) had a high-strong overall rating. Downgrading of articles was generally caused by reliance on using only self-report at the outcome as with the Coaching Boys Into Men program (Miller et al., 2012), limited generalizability with the Support for Students Exposed to Trauma (Jaycox et al., 2009), and high attrition rates with the PSSA program.

Three of the Trauma programs received an overall rating of high-strong including the Fourth R, KiVa (Salmivalli, Kärnä, & Poskiparta, 2011) and Shifting Boundaries programs (Taylor, Stein, Mumford, & Woods, 2013). These programs teach students a variety of skills relevant to problem solving, coping ability, and personal safety awareness. The more effective programs had a high-strong overall rating in their encouragement of developing coping and resiliency skills. Table 8 provides summary of studies related to trauma.

**EDs**

The relevance and importance of EDs in Canada was underscored when Parliament tabled their report in 2014 declaring:
<table>
<thead>
<tr>
<th>Reference</th>
<th>Intervention category</th>
<th>Specific intervention</th>
<th>Program appropriate for grades</th>
<th>Effect size</th>
<th>Quality rating</th>
<th>Strength rating</th>
<th>Overall rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ager et al. (2011)</td>
<td>RCT</td>
<td>PSSA</td>
<td>Study and Program Grades: 2-7</td>
<td>$ES_{\text{range}} = 0.34-0.48$</td>
<td>3</td>
<td>2</td>
<td>3.2</td>
</tr>
<tr>
<td>Ball, Kerig, and Rosenbluth (2009)</td>
<td>Quasi-Experimental</td>
<td>Expect Respect</td>
<td>Study and Program Grades: 7-12</td>
<td>Not reported</td>
<td>1</td>
<td>1</td>
<td>1.1</td>
</tr>
<tr>
<td>Brown et al. (2011)</td>
<td>RCT</td>
<td>Steps to Respect</td>
<td>Study and Program Grades: 2-6</td>
<td>$ES_{\text{range}} = 0.12-0.61$</td>
<td>3</td>
<td>1</td>
<td>3.1</td>
</tr>
<tr>
<td>Cooley-Strickland, Griffin, Darney, Otte, and Ko (2011)</td>
<td>RCT</td>
<td>Modified FRIENDS Program</td>
<td>Study and Program Grades: 3-5</td>
<td>$ES_{\text{range}} = 0.11-0.46$</td>
<td>4</td>
<td>1</td>
<td>4.1</td>
</tr>
<tr>
<td>Fazel, Doll, and Stein (2009)</td>
<td>Quasi-Experimental</td>
<td>Mental Health Service Program</td>
<td>Study and Program Grades: Kindergarten-12</td>
<td>$ES_{\text{range}} = 0.07-0.35$</td>
<td>2</td>
<td>1</td>
<td>2.1</td>
</tr>
<tr>
<td>Gelkopf and Berger (2009)</td>
<td>Quasi-Experimental</td>
<td>ERASE-Stress</td>
<td>Study and Program Grades: 7-8</td>
<td>$ES_{\text{range}} = 0.52-0.72$</td>
<td>3</td>
<td>2</td>
<td>3.2</td>
</tr>
<tr>
<td>Jaycox et al. (2009)</td>
<td>RCT</td>
<td>Support for Students Exposed to Trauma</td>
<td>Study and Program Grades: 6-7</td>
<td>$ES_{\text{range}} = 0.07-0.28$</td>
<td>3</td>
<td>1</td>
<td>3.1</td>
</tr>
<tr>
<td>Salmivalli, Kärnä, and Poskiparta (2011)</td>
<td>RCT</td>
<td>The KiVa Program</td>
<td>Study and Program Grades: 4-6</td>
<td>$ES_{\text{range}} = 0.89-1.97$</td>
<td>4</td>
<td>2</td>
<td>4.2</td>
</tr>
<tr>
<td>Miller et al. (2012)</td>
<td>Quasi-Experimental</td>
<td>Coaching Boys into Men</td>
<td>Study and Program Grades: 9-12</td>
<td>$ES_{\text{range}} = 0.01-0.38$</td>
<td>2</td>
<td>1</td>
<td>2.1</td>
</tr>
<tr>
<td>Taylor, Stein, Mumford, and Woods (2013)</td>
<td>RCT</td>
<td>Shifting Boundaries</td>
<td>Study and Program Grades: 6-7</td>
<td>$ES_{\text{range}} = 0.49-2.07$</td>
<td>4</td>
<td>2</td>
<td>4.2</td>
</tr>
</tbody>
</table>

*Note.* Quality scores range from 1 (weak) to 4 (strong) and strength 1 = weak or 2 = strong. RCT = randomized clinical trial; PSSA = Psychological Structured Activities Program.
At any given time in Canada, as many as 600,000 to 990,000 Canadians may meet the diagnostic criteria for an eating disorder, primarily anorexia nervosa, bulimia nervosa, or binge eating disorder. Approximately 80% of individuals with eating disorders are girls or women. Eating disorders are a serious form of mental illness, “characterized by a persistent disturbance of eating or eating-related behaviour that results in the altered consumption or absorption of food and that significantly impairs physical health or psychosocial functioning.” The devastating symptoms of an eating disorder lead to serious consequences: an individual’s mental and physical health are compromised, personal relationships may suffer, current and future education and employment opportunities may be jeopardized, financial security is put at risk, and overall quality of life deteriorates. (p. 1)

More alarming is that recent Canadian data suggest ED is on the rise and currently the most chronic mental health condition for girls in Canada (Hopewell, 2015). The British Columbia Canadian Mental Health Association (CMHA; 2015) indicated that up to 40% of 9-year-olds have dieted despite being of normal weight. The Canadian National Initiative for Eating Disorders (2015) indicated that while there is high level concern for the 8% of girls who are overweight, the incidence of ED in girls of similar age is actually twice that rate at 18%.

The EDs studies were comprised of seven RCTs and three quasi-experimental studies, ranging from low to high. Sample sizes ranged from 88 to 3,086 participants with a mean age of 12.9 years ($SD = 3.98$). The average program duration was 10.7 weeks ($SD = 7.56$). Six out of 10 studies demonstrated a moderate-to-strong positive overall program effectiveness. Two studies (Stice, Rohde, Shaw, & Gau, 2011; Stice, Shaw, Burton, & Wade, 2006), though very well designed (i.e., high quality), were given a weak recommendation due to their smaller effect sizes.

The ED interventions ranged from short term (3 weeks) to longer term (6 months), with most (60%) being implemented with sessions that occurred once per week. Some of these programs were more intensive, ranging from twice a week (Manios, Moschandreas, Hatzis, & Kafatos, 2002) to 5 times a week (Neumark-Sztainer et al., 2010). Half of these studies did not report information on program training, thus not allowing for a determination of intervention transferability. Thirty percent of the interventions (Kain et al., 2004; Neumark-Sztainer et al., 2010; Stice et al., 2006) involved moderately intensive training requiring attendance at a workshop of 2 days or less. Twenty percent of the interventions (Manios et al., 2002; Stice et al., 2011) involved little to no training reflecting high levels of transferability.

Interventions in this category focused on two different types of EDs: those that resulted in underweight students (e.g., anorexia, bulimia) and those that resulted in overweight students (e.g., overeating, obesity). The most effective interventions for the first type of ED (e.g., anorexia, bulimia) were interactive and psychoeducational in nature. These often involved a lecture component, group discussion, and activities, while some even involved critical analysis of media portrayals to determine their influence on disordered eating (e.g., González, Penelo, Gutiérrez, & Raich, 2011). The most effective interventions for the second type of ED (e.g., overeating, obesity)
targeted multiple factors such as in the New Moves intervention (Neumark-Sztainer et al., 2010) that educated students about a proper diet and nutrition, while also integrating a physical activity component to incorporate exercise that can boost physical health combating a sedentary lifestyle. Finally, the most effective interventions for both types of EDs used a homework component to ensure that what the students learned was being practiced and applied outside of the classroom. This is typically where parents could become involved as well, ensuring that any necessary changes generalized to home (e.g., improving child’s diet, decreasing sedentary activities, etc.) Table 9 provides a summary of studies related to eating disorders.

Discussion

The strength of a literature review is only as good as the existing studies on which it is based. The area of SBMH is fortunate to have a literature that is mature to the point where GRADE analysis is possible across a number of relevant mental health areas. For the current analysis, 98 studies published over the past 10 years were identified that met the criteria of being a two-group comparison of classroom-based interventions with relevant outcome measures. Notably of this number, 67% were RCTs.

Program Effectiveness

Certain of the areas that were reviewed revealed particularly strong overall positive outcomes. This was evident in the addictions area, where half of the programs generated effect size placing them in the highly effective category. To a lesser extent but still extremely encouraging were effect size in the moderate category in the areas of EI, mindfulness, trauma, anxiety, and depression. Where analysis showed particularly poor outcomes were related to suicide prevention programs.

However, beyond the general outcomes across programs, it is encouraging to note that there are programs with effect size in seven of the nine areas considered to be highly effective. These would include those summarized in Table 10.

The two areas where there were no programs considered in the highly effective range were mindfulness and EDs.

Common Program Components

The most effective programs, regardless of their targeted focus of concern, were characterized by common components.

First, effective programs provided sessions over a period of weeks if not months. Programs that were delivered in a single session such as those in the SR area where a mental health survivor would present within a forum generating some dialogue through small group discussion or question and answer were not rewarded with particularly strong positive outcomes. Seven of the 10 SR programs that were part of this evaluation were single session programs. Among the exceptions to this was the Understanding Our Peers initiative (Ranson & Byrne, 2014) that, while providing exposure to a
Table 9. Summary of Studies Related to Eating Disorders.

<table>
<thead>
<tr>
<th>Reference</th>
<th>Intervention category</th>
<th>Specific intervention</th>
<th>Age range</th>
<th>Effect size</th>
<th>Quality rating</th>
<th>Strength rating</th>
<th>Overall rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stice, Shaw, Burton, and Wade (2006)</td>
<td>RCT</td>
<td>Dissonance and Healthy Eating Programs</td>
<td>14-19 years 9th-12th grade</td>
<td>$r_{range} = .11-.38$ $r_{mean} = .20$</td>
<td>4</td>
<td>1</td>
<td>4.1</td>
</tr>
<tr>
<td>Stice, Rohde, Shaw, and Gau (2011)</td>
<td>RCT</td>
<td>Dissonance Program</td>
<td>14-19 years 9th-12th grade</td>
<td>$E_{range} = 0.15-0.39$ $E_{mean} = 0.26$</td>
<td>4</td>
<td>1</td>
<td>4.1</td>
</tr>
<tr>
<td>González, Penelo, Gutiérrez, and Raich (2011)</td>
<td>Quasi</td>
<td>Eating, Aesthetic Feminine Models and the Media</td>
<td>12-14 years 7th-8th grade</td>
<td>$d_{range} = 0.29-0.78$ $d_{mean} = 0.64$ $r_{range} = .19-.36$ $r_{mean} = .30$</td>
<td>2</td>
<td>2</td>
<td>2.2</td>
</tr>
<tr>
<td>Bird, Halliwell, Diedrichs, and Harcourt (2013)</td>
<td>Quasi</td>
<td>Happy Being Me</td>
<td>10-11 years 5th grade</td>
<td>$d_{range} = 0.30-0.92$ $d_{mean} = 0.65$</td>
<td>2</td>
<td>2</td>
<td>2.2</td>
</tr>
<tr>
<td>Kain et al. (2004)</td>
<td>Quasi</td>
<td>Healthy School</td>
<td>6-14 years 1st-8th grade</td>
<td>$d_{range} = 0.21-0.55$ $d_{mean} = 0.42$ $r_{range} = .10-.26$ $r_{mean} = .20$</td>
<td>2</td>
<td>2</td>
<td>2.2</td>
</tr>
<tr>
<td>Fitzgibbon et al. (2005)</td>
<td>Modified RCT</td>
<td>Hip Hop to Health Jr.</td>
<td>4-5 years Kindergarten</td>
<td>$d_{range} = 0.15-4.71$ $d_{mean} = 1.96$ $r_{range} = .08-.92$ $r_{mean} = .57$</td>
<td>3</td>
<td>2</td>
<td>3.2</td>
</tr>
<tr>
<td>Neumark-Sztainer et al. (2010)</td>
<td>Modified RCT</td>
<td>New Moves</td>
<td>14-19 years 9th-12th grade</td>
<td>$d_{range} = 0.18-8.66$ $d_{mean} = 1.20$ $r_{range} = .09-.97$ $r_{mean} = .35$</td>
<td>3</td>
<td>2</td>
<td>3.2</td>
</tr>
<tr>
<td>M. Jones et al. (2008)</td>
<td>RCT</td>
<td>Student Bodies</td>
<td>14-19 years 9th-12th grade</td>
<td>$E_{range} = 0.54-1.20$ $E_{mean} = 0.79$</td>
<td>3</td>
<td>2</td>
<td>3.2</td>
</tr>
<tr>
<td>Manios, Moschandreas, Hatzis, and Kafatos (2002)</td>
<td>RCT</td>
<td>Know Your Body</td>
<td>6-12 years 1st-6th grade</td>
<td>$d_{range} = 0.15-0.62$ $d_{mean} = 0.35$ $r_{range} = .07-.30$ $r_{mean} = .17$</td>
<td>4</td>
<td>2</td>
<td>4.2</td>
</tr>
<tr>
<td>D. A. Stewart, Carter, Drinkwater, Hainsworth, and Fairburn (2001)</td>
<td>RCT</td>
<td>Unnamed Eating Disorder Prevention Program</td>
<td>14-15 years 9th grade</td>
<td>$d_{range} = 0.32-0.72$ $d_{mean} = 0.58$ $r_{range} = .16-.34$ $r_{mean} = .28$</td>
<td>2</td>
<td>2</td>
<td>2.2</td>
</tr>
</tbody>
</table>

Note. Quality scores range from 1 (weak) to 4 (strong) and strength 1 = weak or 2 = strong. RCT = randomized clinical trial.
mental health survivor, also involved an ongoing homework component where participants were asked to complete online activities on a web-based learning site.

Second, more effective programs required at least some training and ongoing supervision to insure program fidelity. Although this added to the cost and investment in time by the teacher, extended training, characterized for this review as compromising 2 days or more, along with ongoing support from someone attached to the program designer, contributed to stronger outcomes.

Third, effective programs drew on parent support or required some form of homework beyond the classroom to insure a degree of generalizability.

Fourth, effective programs were manualized such that teachers could reference the particular curriculum component to insure continuity and consistency with the intent of the program.

Beyond Effectiveness

Beyond the effectiveness of programs, the GRADE analysis provides an additional means by which studies can be evaluated to address the issues of application and transferability.

Schools need to acknowledge more than the extent to which programs are effective. For an intervention to be integrated within the culture of a school, it is also important

---

Table 10. Programs Receiving a Rating of Highly Effective Based on Overall Outcome Effect Sizes.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Programs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Emotional Intelligence</td>
<td>PATHS Project (Shek &amp; Ma, 2012)</td>
</tr>
<tr>
<td></td>
<td>RULER Approach (Bond et al., 2004)</td>
</tr>
<tr>
<td></td>
<td>Fourth R (Crooks, Scott, Ellis, &amp; Wolfe, 2011)</td>
</tr>
<tr>
<td>Stigma Anxiety</td>
<td>Stigma Reduction Program (Saporito et al., 2011)</td>
</tr>
<tr>
<td>Anxiety</td>
<td>Confident Kids (Berry &amp; Hunt, 2009)</td>
</tr>
<tr>
<td>Depression</td>
<td>Coping with Stress Course (Horowitz et al., 2007)</td>
</tr>
<tr>
<td></td>
<td>LARS &amp; LISA (Pössel et al., 2008)</td>
</tr>
<tr>
<td>Addictions</td>
<td>Positive Action (Beets et al., 2009)</td>
</tr>
<tr>
<td></td>
<td>CHOICE (D’Amico et al., 2012)</td>
</tr>
<tr>
<td></td>
<td>Unplugged (Faggiano et al., 2010)</td>
</tr>
<tr>
<td></td>
<td>Media Detective (Kupersmidt et al., 2010)</td>
</tr>
<tr>
<td></td>
<td>European Drug Abuse Program (Vigna-Taglianti et al., 2009)</td>
</tr>
<tr>
<td>Suicide Prevention</td>
<td>South Elgin High School Suicide Prevention Program (Ciffone, 2007)</td>
</tr>
<tr>
<td>Trauma</td>
<td>The KiVa Program (Salmivalli et al., 2011)</td>
</tr>
</tbody>
</table>

Note. RULER = Recognizing, Understanding, Labeling, Expressing, and Regulating.

*LARS & LISA are gender aware programs that promote understanding of the key concepts through exploring the relationship between cognition, emotion, and behavior.
for it to be accessible and transferable. For teachers, accessibility means an intervention has to be available at low/no cost with a reasonable commitment to the training and ongoing supervision along with the amount of class time required to deliver the intervention. In these contexts, the GRADE analysis provides ratings beyond effectiveness that focus on the strength and quality of implementation. Downgrading of the quality of studies reflected the nature of the design, for example, a quasi-experimental rather than an RCT, lack of a fidelity measure, and/or a smaller sample size, effect size outcomes, and sample attrition rate. It also included a judgment on the degree of practicality for implementation that accounted for a program’s cost, length of intervention and/or training, and whether it was manualized.

For program selection based on a GRADE analysis, school administrators and teachers need to select from programs that have a track record of success, that are effective with the targeted age and grade of students, and that are feasible in cost and implementation.

**Next Steps**

The value of a GRADE review lies in its support for decision making in regard to the application of evidenced informed practice for SBMH interventions. The current GRADE analysis reflected general outcomes related to the interventions under review. Future reviews will want to focus more specifically on the types of outcomes being reported. For example, many of the programs mentioned in this review were relatively brief in duration, and expectations that there would be significant shifts in cognitions or attitudes related to certain interventions may not have been possible with the follow-up periods that were being reported. In addition, as the number of studies in each area increases, future reviewers will want to begin to differentially weight the outcomes from intervention based on the sample sizes within each study in addressing the potential biases inherent in averaging effect sizes.

Our next steps with respect to the findings from this review include the creation of an accessible means by which educators can draw on these findings to facilitate the informed selection of appropriate programs that fit with the grade level, time allocation, requirements for training, and cost. This will be facilitated through the development of a free and accessible online accessible program that leads educators through the program components of each focused area. The development of such an accessible system to support teachers is currently underway by the authors along with technological support from the Portico working group at the Centre for Addictions and Mental Health.

**Summary**

This GRADE analysis focused on nine areas of concern to teachers regarding the well-being of their students. An analysis based on effectiveness, quality, and strength provided a profile of 98 programs that met the eligibility criteria for inclusion in the review. The summary of these programs provides yet another means by which school personnel can base their decisions when considering the adoption of a SBMH strategy for their school.
Declaration of Conflicting Interests
The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding
The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported through a grant from Physical and Health Education Canada.

References
Ball, B., Kerig, P. K., & Rosenbluth, B. (2009). “Like a family but better because you can actually trust each other”: The expect respect dating violence prevention program for at-risk youth. Health Promotion Practice, 10, 45S-58S.


Cunningham, E. G., Brandon, C. M., & Frydenberg, E. (2002). Enhancing coping resources in early adolescence through a school-based program teaching optimistic thinking skills. Anxiety, Stress & Coping, 15, 369-381.


**Author Biographies**

**Susan Rodger** is an associate professor in the Faculty of Education at Western University. Dr. Rodger’s research interests include school based emotional health for teachers and students.

**Renelle Bourdage** is a second year MA student in Western’s Graduate Program in Counselling Psychology. Renelle studies mental health supports for teachers.
Kaitlin Hancock is a graduate of Western University’s Counselling Psychology Program. Kaitlin is a therapist with expertise in the measurement of mental health.

Rebecca Hsiang is a graduate of Western University’s Counselling Psychology Program. Rebecca is a therapist whose research interests include emotional intelligence and mental health.

Robyn Masters is in her second year of doctoral studies in Western University’s School and Applied Child Psychology Program, where she studies teacher education to promote mental health literacy.

Alan Leschied is a professor in the Faculty of Education at Western University. Dr. Leschied’s research interests are in areas related to children’s mental health, child welfare and the assessment of youth involved in the justice system.