REVIEW PROTOCOL FOR STUDIES OF PREVENTIVE INTERVENTIONS TO SUPPORT STUDENT MENTAL HEALTH AND WELL-BEING, VERSION 3.0

TOPIC AREA FOCUS

This protocol guides the review of research that informs What Works Clearinghouse (WWC) intervention reports in the area of student mental health and well-being. This review-specific protocol is used in conjunction with the *WWC Procedures and Standards Handbook (Version 3.0)*.

This review focuses on interventions for middle, high school, and postsecondary students that aim to improve mental health and well-being and reduce externalizing problems, internalizing problems, and substance use.

Mental health and well-being can be conceptualized as including psychological and emotional states and processes commonly associated with health and wellness, such as happiness, self-esteem, feelings of connectedness, and life satisfaction, as well as those regarded widely as important for personal growth and achievement, such as a sense of competence, intrinsic motivation, and capacity for independent action (Department of Health and Human Services, 1999). The mental health and well-being of students are important from both a public health perspective (e.g., because initial onset for many mental health disorders occurs during adolescence) and an educational perspective (i.e., because well-functioning students are more likely to perform to their peak level academically). As students in this age range (roughly 12-22) navigate stressors in multiple domains -- physiologically, socially, and academically -- ineffective or limited coping strategies may contribute to mental health problems (Ryan & Deci, 2001).

The prevalence of mental health and well-being problems among students in this age range has been well-documented. Recent large-scale survey research has indicated that nearly half (49%) of community college students show symptoms related to one or more mental health condition, such as depression, anxiety, suicidal ideas, nonsuicidal self-injury, or eating disorders (Eisenberg, Goldrick-Rab, Lipson, & Broton, 2016). A similar trend was found through a survey of four-year universities: nearly one-third (32%) of students indicated that they have suffered from depression, anxiety, or suicidal thoughts in the past. Nearly half of those students (45%) said that mental health problems have affected their academic performance in the past 1-5 days (Eisenberg, Hunt, & Speer, 2013). Adolescent students share similar problems as well. Results from the National Comorbidity Survey - Adolescent Supplement indicated that more than 1 in 5 students suffer from a DSM-IV mental disorder. The most common problems students reported were anxiety (31.9%), behavior, (19.1%), or mood disorders (14.3%) (Merikangas et al., 2010).

Internalizing problems like depression and anxiety are one aspect of a student's mental health and well-being, and meta-analytic research has also linked internalizing problems to academic performance. We define internalizing problems as an inward state or affect that inhibits daily life functioning either personally or academically. A recent meta-analysis indicated that students who suffer from both general and academic stress tend to have lower grade-point averages (Richardson, Abraham, & Bond, 2012). Furthermore, a meta-analysis of longitudinal studies found that students with more emotional problems, such as anxiety, depression, or general internalizing symptoms, were more likely to experience negative educational outcomes like dropping out of school or failing to earn a postsecondary credential (Riglin, Petrides, Fredrickson, & Rice, 2014).

Externalizing disorders also have the potential to impact a students' well-being. Broadly defined, any action that disrupts the social environment constitutes an externalizing problem. More specifically, externalizing problems can be defined as outward negative actions that conflict with typical or normative behaviors. These include, but are not limited to, impulsivity and/or hyperactivity, bullying, violence, and/or other antisocial behaviors. Preventing these behaviors is important from both public health and educational perspectives. From a public health perspective, recent meta-analytic results revealed that students who are involved in direct or indirect aggressive behaviors tend to suffer from greater internalizing problems (e.g., depression, anxiety), peer rejection, and emotional dysregulation (Card, Stucky, Sawalani, & Little, 2008), and tend to have high rates of suicidal ideation and behaviors (Holt et al., 2015). From an educational perspective, students with externalizing behavior problems in adolescence tend to have academic competence problems later in adolescence and adulthood (Masten et al., 2005).

Finally, this review is also concerned with students' general well-being in a positive sense. As explained above, these include ones' processes commonly associated with health and wellness, such as happiness, self-esteem, feelings of connectedness, and life satisfaction, as well as those regarded widely as important for personal growth and achievement, such as a sense of competence and capacity for independent action. Meta-analytic research has found that constructs such as optimism, self-esteem, and academic self-efficacy are all positively correlated with college GPA (Richardson et al., 2012). These behaviors and processes can be internal or external characteristics.

To address these and other issues, school-based prevention and intervention programs have been proposed, tested, and documented. These interventions are diverse in the sense that they can target different mechanisms (e.g., stress management, attributional style), different mental health issues (e.g., externalizing, substance use) alone or in combination, and employ different levels of intervention (e.g., universal vs. targeted). Many prevention programs enact change through large-scale implementation that allows many or all students within a given school or district, to be targeted at once. These programs can be relatively straightforward for schools to implement because, compared to programs that target interventions to a small number of students already exhibiting problems, they often rely on teachers or in-place staff to implement the programs.

RESEARCH QUESTIONS GUIDING THIS REVIEW

The following questions guide the review and inclusion of studies to improve students' mental health and well-being:

- Does the reviewed intervention appear to be effective, at the middle and high school level, for increasing academic achievement and graduation, reducing dropout from high school, increasing positive internalizing or externalizing behaviors, decreasing negative internalizing or externalizing behaviors, and/or decreasing substance use?
- At the postsecondary level, does the reviewed intervention appear to be effective in increasing postsecondary access and enrollment, enrollment persistence, credit accumulation and attainment, improving academic achievement in college, increasing positive internalizing and externalizing behaviors, decreasing negative internalizing or externalizing behaviors, improving labor market outcomes, and/or decreasing substance use?
- Is the reviewed intervention more or less effective for certain subgroups of students (including first-generation college students, racial/ethnic minorities, academically underprepared students, students from low socioeconomic status backgrounds [e.g., percent of students with free- or reduced-priced lunch], and/or community college students)?

IDENTIFYING STUDIES FOR REVIEW

The *WWC Procedures and Standards Handbook* articulates the general procedures for conducting a literature search. For the mental health and well-being topic area, a broad search will be conducted to identify potentially relevant intervention studies. In addition, for each intervention report under this topic area, a secondary search will be performed to identify any studies of the intervention that were not identified in the initial search (including a search of the What Works Clearinghouse's database of reviewed studies). Further, once interventions have been identified as being targets for an Intervention Report, the WWC supplements the electronic database search with targeted searches of government and non-government agency websites, relevant non-profit organizations that might fund research on the intervention, and via reviewing the bibliographies of literature reviews, meta-analyses, and primary studies of the intervention under review.

The broad search for the student mental health and well-being topic area is detailed in Appendix A. Each Intervention Report's secondary search will be described in Appendix B.

ELIGIBILITY AND INCLUSION CRITERIA

Studies must meet several criteria to be eligible for review under the student mental health and well-being topic area. These relate to the population that was studied, the study design that was used, the outcomes that were measured, and when the study was conducted. Each of these is discussed below.

Population Included

This review protocol will include interventions designed to improve mental health and well-being for students of different ages. Therefore, this review specifies upper and lower age boundaries and then categorizes eligible interventions or strategies according to the age of the students when interventions take place (i.e., middle school, high school, or college).

To be eligible for review under this protocol, a study must include students in middle school, high school, or postsecondary institutions. For the *lower age boundary*, this review includes activities that start at the beginning of middle school/junior high or age 11. The 6th grade is considered the start of middle school/junior high. Studies that focus on students in elementary school, or include students younger than 11 years old, will not be included. If a sample of middle school aged students includes students under the age of 11, the study is eligible for review if the average age of the sample is over 11 or if the 11-14 year olds make up at least 60% of the sample. At the middle and high school level, no upper age boundary is set, but the students included in the sample should constitute "traditional" high school aged students (i.e., most should be 18 and younger).

Because postsecondary institutions enroll a wide range of students, this review applies an *upper age boundary* to students in postsecondary education. This review will focus only on traditional aged students, which are defined as ages 18-24. This period is when the vast majority of students experience their most intense spell of higher education. Older adults enrolled in higher education may have different issues and are likely to be targeted by different sorts of interventions. If a sample of college students includes students over the age of 24, the study is eligible for review if the average age of the sample is under 25 or 18-24 year olds make up at least 60% of the sample. If the sample includes only postsecondary students, it is permissible to include samples that have some students under the age of 18.

In addition, regardless of age, the review will exclude studies with samples that have more than 10% of students with an identified mental health disorder[1]. In addition, studies that include samples with more than 50% of students with an identified learning disability will also be excluded. No other sample defining characteristics (e.g., gifted status) will be used to exclude studies.

In general, the WWC determines a study rating based on average intervention effects and will report subgroup analyses only for groups that are identified in the protocol as being of theoretical, policy, or practical interest. For studies reviewed under this protocol, the default subgroups include: (a) gender, (b) student who are or would be first-generation college students, (c) racial/ethnic minorities, and/or (d) students from low socioeconomic status backgrounds (e.g.,

[1] "Identified mental health disorder" includes official diagnoses of a mental health disorder (e.g., by a mental health professional) and scoring in a pre-established clinical range of a well-established measure.

Pell Grant recipients, or students eligible for free or reduced price lunch). To be eligible for review as a subgroup analysis, impact estimates must be available for all groups in a subgroup analysis (e.g., results for both first-generation college students *and* non-first-generation college students, not just one or the other) and a test of the interaction between subgroup membership and intervention condition must be reported or derivable from reported statistics (using, for example, techniques described in Altman & Bland, 2003).

As discussed in the *WWC Procedures and Standards Handbook* (v. 3.0, see Section III.B.4, p. 17), if a study presents findings separately for several groups of students without presenting an aggregate result, the WWC will query authors to determine if they conducted an analysis on the full sample of students. If the WWC is unable to obtain aggregate results from the author, the WWC averages across subgroups within a study to use as the primary finding and presents the subgroup results as supplemental analyses.

Interventions Included

Scope

Studies eligible for review under this protocol must be school-based and focus on students' mental health and well-being, broadly defined.

School-based interventions are those that are implemented in a middle school, high school, or postsecondary institutional setting. Programs that occur outside a school setting, for example at a community center, are not eligible for review. For postsecondary programs, only programs implemented by the institution's staff (i.e., individuals employed by or currently enrolled at the institution) will be included. Studies of interventions that are implemented by postsecondary staff in conjunction with individuals not employed by the postsecondary institution are eligible for review.

School-based preventive interventions use a variety of methods to promote student mental health and well-being. These interventions can be classified into two broad types: universal prevention interventions and targeted prevention interventions. Two features distinguish universal prevention programs. First, in universal prevention programs the intervention is delivered to an intact group of students (e.g., every student in a seventh grade class; every seventh grade class in a school; every student in a middle school). In addition, the students receiving the intervention have not been individually identified as needing the preventive intervention.

Targeted prevention programs individually identify students for intervention. Generally, these students will exhibit sub-clinical signs of a disorder (e.g., might report multiple symptoms of depression but not at the level that would warrant a diagnosis of depression) and the intervention will be delivered in small groups (or perhaps individually).

Finally, some prevention interventions have features of both universal and targeted prevention, in that they are delivered to entire groups of students that researchers identify as being 'at-risk' given their location in a specific school, community, or context.

All of the above models are eligible for review under this protocol. Programs that specifically target students with an identified clinical disorder are not eligible.

Content

To be eligible for review, the content of the intervention must focus on students' mental health and well-being. Examples of content include programs that prevent or decrease students' test anxiety or other related anxieties related to attending school; (b) programs that attempt to improve students' self-efficacy, resiliency, or self-worth; (c) targeted substance use and abuse programs that attempt to decrease students' use of alcohol or drugs. Programs that focus on specific aspects of performing in or attending school, such as targeted dropout prevention programs or dedicated literacy coaching, are not eligible for review under this protocol.

As the literature search is completed, this protocol may be updated to specify Intervention Reports or products that include specific criteria. To date, this includes:

- **Interventions to reduce stress for students preparing for college.** As students progress through middle and high school, they are asked to complete various college preparation tasks, for example, college preparation courses, college entrance examinations, exit exams, and financial aid applications. These tasks may induce undue stress and anxiety that impacts students' ability to function productively. The purpose of this Intervention Report, therefore, will be to review interventions that target these problem behaviors.

Location

All programs must be implemented in the United States or Canada. Studies that include programs from outside the U.S. or Canada will not be included unless the analysis disaggregates the sample by country.

Types of Studies

To be eligible for review, a study must be a primary analysis of the effects of an intervention. If a study does not examine the effects of an intervention, or if it is not a primary analysis (e.g., if it is a meta-analysis or other literature review), it is not eligible for review.

In addition, the study must have an eligible design. Eligible study designs include randomized controlled trials and well-controlled quasi-experimental designs (defined as studies using a well-matched comparison group). The WWC currently does not have standards for some other types of quasi-experimental designs, such as the instrumental variable approach. In addition, studies using regression discontinuity designs or single-case designs will be eligible under this topic

area; should any studies with these designs be located for review, the pilot standards described in the *WWC Procedures and Standards Handbook* will be used.

Types of Comparisons

Studies reviewed under this protocol for Intervention Reports must use either "business as usual" or a "no treatment" comparison group. "Business as usual" comparison groups are those in which students may attend the same or similar schools as the intervention students and/or they may receive the usual services offered to students in the setting (e.g., advising, tutoring). Comparison groups must not involve explicit assignment of students to other putatively effective interventions or variations of the same intervention that is delivered to the intervention group. These studies are commonly referred to as "comparative effectiveness research" or "multiple arm treatment research." Furthermore, should studies include multiple intervention groups and an included comparison group (i.e., business-as-usual), only the comparisons between the intervention and comparison group will be included; the comparisons between intervention groups will be excluded. Studies for which the type or nature of the comparison group is not clearly "business as usual" should be referred to the review team leadership for consultation, to ensure that comparison conditions are similar across studies.

Studies to be reviewed for Quick Reviews and Single Study Reviews (including those reviewed for U.S. Department of Education funding competitions) may include comparison groups that receive other or similar interventions as well as "business as usual" comparison groups.

Eligible Outcomes

To be eligible for review, a study must assess a relevant outcome domain. These may include outcomes measured prior to attending a postsecondary institution while students are in middle or high school or outcomes assessed while students are transitioning to or attending postsecondary institutions. The following middle or high school outcome domains are eligible: (a) academic achievement, (b) attendance, (c) staying in school, (d) progressing in school, (e) completing school, (f) postsecondary enrollment, (g) internalizing behaviors, (h) externalizing behaviors, and/or (i) substance use. The following postsecondary outcome domains are eligible: (a) enrollment, (b) attendance, (c) credit accumulation and persistence, (d) academic achievement, (e) attainment, and (f) internalizing behaviors, (g) externalizing behaviors, (h) substance use/abuse, and/or (i) labor market outcomes. Operational definitions for each outcome domain are provided below.

Finally, authors may report measures from various data collection time points. For all middle and high school outcomes, for postsecondary enrollment, postsecondary attendance, and for postsecondary measures of internalizing and externalizing outcomes, reviews will focus on the most immediate measure after the intervention. For postsecondary credit accumulation, postsecondary academic achievement, postsecondary attainment, and postsecondary labor market outcomes, reviews will focus on the longest follow-up time point.

The content expert is responsible for grouping outcomes into domains. These are defined as follows:

Middle and High School Outcomes

- **Academic Outcomes**
  - **Academic achievement** in middle or high school refers to the extent to which students adequately complete expected coursework. Examples of ways that academic achievement might be operationally defined in studies include: (a) standardized achievement tests including the ACT and SAT as well as state-mandated tests, (b) high school grade point averages, or (c) AP test results. Individual course grades and exam scores from middle or high school courses are not eligible under this domain.
  - **Attendance in middle or high school** refers to outcomes that measure attendance rates or absenteeism at school. Ways that attendance might be operationalized include the number or proportion of days absent or in attendance during a school term, proportion of students with excessive absences, and referrals for truancy. Objective measures of attendance, such as those from school administrative records are preferred, but student reported measures are acceptable if a more objective measure is not available.
  - **Staying in school** refers to outcomes that measure whether the student has dropped out of school and the number of days the student was enrolled in school.
  - **Progressing in school** refers to outcomes that assess the number of high school course credits the student has earned, whether the student was promoted to the next grade, and the highest grade the student has completed.
  - **Completing school** refers to outcomes that measure whether the student has earned a high school diploma or GED.
  - **Labor market outcomes** refer to outcomes related to employment after the postsecondary experience. Examples of ways that labor market outcomes might be operationally defined in studies include (a) employed vs. not, (b) employed full-time vs. employed part-time, (c) employed in field of study vs. not, and (d) income earned.

- **Internalizing and Externalizing Outcomes**
  - **Internalizing behaviors or symptoms** includes two aspects of an individual's well-being and encompass an individual's thoughts, experiences, and processes. One aspect, *negative* internalizing behaviors include, but are not limited to, depression and depressive symptoms; suicidality; somatic complaints; phobia, panic, and anxiety disorders and symptoms; and loneliness or despair. The other aspect, *positive* internal behaviors and symptoms include happiness, positive well-being, general life satisfaction, or self-esteem. Either type of internal behavior or symptom, negative or positive, is eligible for review.

- **Externalizing behaviors or symptoms** include those that are directed outward. Externalizing behaviors can manifest through disruptive social interactions, such as through conduct disorders, malicious actions, or violent behavior, or through prosocial behaviors, for example developing healthy friendships, tutoring peers, willingness to help other people, or intervening to prevent bullying. Either type of external behavior, negative or positive, is eligible for review.

Postsecondary Outcomes

- **Academic Outcomes**
  - **Access and enrollment** refers to the process of applying to, actually enrolling, and attending a postsecondary institution. Examples of ways that enrollment might be operationally defined in studies include: (a) actual enrollment in college; (b) number and/or selectivity of admitted and/or enrolling institutions, (c) enrollment by institution type (2 year vs. 4 year), (d) intensity of enrollment (full time vs. part time), and (e) timing of enrollment (e.g., immediate vs. delayed enrollment after high school). On a case-by-case basis, the WWC's postsecondary content expert may accept measures of intentions to enroll, though if measures of actual enrollment are preferred when both types are available.
  - **College attendance** refers to outcomes that measure attendance rates or absenteeism. Ways that attendance might be operationalized include the number or proportion of days absent or in attendance during a school term, proportion of students with excessive absences, and the like. Objective measures of attendance, such as those from school administrative records are preferred, but student reported measures are acceptable if a more objective measure is not available.
  - **Credit accumulation and persistence** refers to progress toward the completion of a degree, certificate, or program. Examples of ways that credit accumulation might be operationally defined in studies include: (a) number of college-level credits earned, (b) number of terms of continuous enrollment, and (c) enrolled vs. did not enroll the next semester. The number of non-college level credits earned (e.g., developmental credits) is not an eligible measure of credit accumulation. On a case-by-case basis, the WWC's postsecondary content expert may accept measures of intentions to persist, though measures of actual persistence are preferred when both types are available.
  - **Academic achievement** refers to the extent to which students adequately complete expected coursework. As such, eligible measures of academic achievement are those that arise naturally from student educational experiences. Examples of ways that academic achievement might be operationally defined in studies include (a) final grade in a single college-level course, (b) grade point average in college-level courses, and (c) the ratio of college-level courses passed vs. failed. Scores on professional or industry exams (e.g., the GRE and the NCLEX-RN) are also eligible. With the exception of department-wide final exams, measures that exist below the final course grade level are not eligible (e.g.,

average test score, score on a particular assignment or project). Also ineligible are measures of academic achievement that do not directly contribute to student grades (e.g., a math test that is given after an experimental manipulation, the performance on which has no implications for a student's performance in a specific course).
    - **Attainment** refers to the completion of a degree, certificate, or program. Examples of ways attainment might be operationally defined in a study include (a) certificate completion rates and (b) degree completion rates.
    - **Labor market outcomes** refer to outcomes related to employment after the postsecondary experience. Examples of ways that labor market outcomes might be operationally defined in studies include (a) employed vs. not, (b) employed full-time vs. employed part-time, (c) employed in field of study vs. not, and (d) income earned.

- **Internalizing and Externalizing Outcomes**
    - **Internalizing behaviors or symptoms** includes two aspects of an individual's well-being and encompass an individual's thoughts, experiences, and processes. One aspect, *negative* internalizing behaviors include, but are not limited to, depression and depressive symptoms; suicidality; somatic complaints; phobia, panic, and anxiety disorders and symptoms; and loneliness or despair. The other aspect, *positive* internal behaviors and symptoms include happiness, positive well-being, general life satisfaction, or self-esteem. Either type of internal behavior or symptom, negative or positive, is eligible for review.
    - **Externalizing behaviors or symptoms** include those that are directed outward. Externalizing behaviors can manifest through disruptive social interactions, such as through conduct disorders, malicious actions, or violent behavior, or through prosocial behaviors, for example developing healthy friendships, tutoring peers, willingness to help other people, or intervening to prevent bullying. Either type of external behavior, negative or positive, is eligible for review.

Timeframe for Studies

Studies must have been published or reported in 1996 or later to be eligible for review under this protocol.

REVIEW OF STUDIES AGAINST WWC STANDARDS

All studies will be reviewed against the WWC Evidence Standards, using version 3.0 of the *Procedures and Standards Handbook.* Generally, these standards assess outcome reliability and validity, attrition, baseline equivalence, and similar methodological and statistical issues. This review determines the overall WWC study rating (see the *Procedures and Standards Handbook Version 3.0* for further details). Details relate to sample attrition in RCTs and baseline equivalence in QEDs and high-attrition RCTs are further articulated in this protocol.

Sample Attrition

The *WWC Procedures and Standards Handbook* discusses the sample attrition standards used by the WWC.

This review uses the liberal boundary for attrition. The selection of this boundary was based on the assumption that most attrition in studies of interventions focused on mental health and well-being is due to factors that are not strongly related to intervention status.

Baseline Equivalence

If the study design is a randomized controlled trial or regression discontinuity design with high levels of attrition, or a quasi-experimental design, the study must demonstrate baseline equivalence of the intervention and comparison groups for the analytic sample.

If demonstration of baseline equivalence is required for a study, the WWC will assess equivalence on a pre-intervention measure of the outcome (i.e., a pretest) or a close proxy.

When pretests or close proxies are not available, studies must demonstrate baseline equivalence on the following:

If the outcome is a measure of academic achievement, a continuously-scaled baseline measure of academic achievement and one of the following (in order of preference)

- Student socioeconomic status (e.g., FAFSA expected family contribution, family income, free- or reduced-price lunch status, parent education levels, Pell grant eligibility).
- Student age or grade level. Age is preferred. Grade level can be matched by design (i.e., if no specific age information is available, and all students are in the 3rd grade, then students are considered to be matched on this dimension).
- Student gender.
- Student race/ethnicity.

If the outcome is a measure of mental health/well-being, two of the following (in order of preference)

- Student academic achievement (note: in this case, non-continuous measures of academic achievement, e.g., special education placement, are acceptable).
- Student socioeconomic status (e.g., FAFSA expected family contribution, family income, free- or reduced-price lunch status, parent education levels, Pell grant eligibility).
- Student age or grade level. Age is preferred. Grade level can be matched by design (i.e., if no specific age information is available, and all students are in the 3rd grade, then students are considered to be matched on this dimension).

- Student gender.
- Student race/ethnicity.

In cases where multiple baseline measures of a characteristic are available, the content expert is responsible for selecting the variable(s) to be used in the baseline equivalence assessment prior to the equivalence assessment being performed. For example, if both math and verbal scores on a college entrance exam are available, and the primary outcome is whether or not students passed their first college level math course, then the content expert may decide that the score on the math portion of the entrance exam is the only achievement measure on which baseline equivalence should be assessed. However, if the primary outcome is attainment, then the content expert might decide to assess balance on both the math subtest and the verbal subtest.

Procedures for Statistical Adjustment for Studies with Baseline Covariate Imbalance

These procedures apply to all studies for which baseline equivalence must be demonstrated (i.e., RCTs with high attrition and quasi-experimental studies)

If a pretest is available for an outcome and the difference between conditions at baseline is shown to be within the range that requires statistical adjustment, the statistical adjustment is only needed for that outcome. For example, if vocabulary, reading comprehension, and reading fluency are available as pre- and post-intervention measures, and the pre-intervention difference in reading comprehension requires statistical adjustment, only the analysis of reading comprehension must adjust for baseline differences in reading comprehension.

For outcomes that do not have a pretest or close proxy, if the difference between conditions at baseline on one of the required covariates is shown to be within the range that requires statistical adjustment, the adjustment is required only for the covariate in the adjustment range. For example, if academic achievement is judged to be within the range that requires statistical adjustment and SES is very closely balanced (i.e., less than or equal to .05 standard deviations separate the means of the intervention and comparison groups), then all outcomes without pretests must adjust for the measure of academic achievement, and adjustment for baseline SES is not required.

REFERENCES

Altman, D. G., & Bland, J. M. (2003). Interaction revisited: The difference between two estimates. *British Medical Journal, 326* (7382)*,* 219.

Beyers, J. M., Toumbourou, J. W., Catalano, R. F., Arthur, M. W., & Hawkins, J. D. (2004). A cross-national comparison of risk and protective factors for adolescent substance use: The United States and Australia. *Journal of Adolescent Health*, *35*(1), 3-16.

Card, N. A., Stucky, B. D., Sawalani, G. M., & Little, T. D. (2008). Direct and indirect aggression during childhood and adolescence: A meta‐analytic review of gender differences, intercorrelations, and relations to maladjustment. *Child development*, *79*(5), 1185-1229.

Fox., L., Carta, J., Strain, P., Dunlap, G., & Hemmeter, M.L. (2009). Response to intervention and the pyramid model. Tampa, FL: University of South Florida, Technical Assistance Center on Social Emotional Intervention for Young Children.

Department of Health and Human Services. (1999). *Mental health: A report of the Surgeon General*. Retrieved from: https://profiles.nlm.nih.gov/ps/retrieve/ResourceMetadata/NNBBHS

Holt, M. K., Vivolo-Kantor, A. M., Polanin, J. R., Holland, K. M., DeGue, S., Matjasko, J. L., ... & Reid, G. (2015). Bullying and suicidal ideation and behaviors: a meta-analysis. *Pediatrics*, *135*(2), e496-e509.

Kentucky Department of Education. (2014). *Suicide prevention and awareness*. Retrieved from: http://education.ky.gov/school/sdfs/pages/suicide-prevention-and-awareness.aspx

Kessler, R. C., McLaughlin, K. A., Green, J. G., Gruber, M. J., Sampson, N. A., Zaslavsky, A. M., ... & Benjet, C. (2010). Childhood adversities and adult psychopathology in the WHO World Mental Health Surveys. *The British Journal of Psychiatry*, *197*(5), 378-385.

Moore, T. H., Zammit, S., Lingford-Hughes, A., Barnes, T. R., Jones, P. B., Burke, M., & Lewis, G. (2007). Cannabis use and risk of psychotic or affective mental health outcomes: A systematic review. *The Lancet*, *370*(9584), 319-328.

Poulton, R., Caspi, A., Milne, B. J., Thomson, W. M., Taylor, A., Sears, M. R., & Moffitt, T. E. (2002). Association between children's experience of socioeconomic disadvantage and adult health: a life-course study. *The Lancet*, *360*(9346), 1640-1645.

Richardson, M., Abraham, C., & Bond, R. (2012). Psychological correlates of university students' academic performance: a systematic review and meta-analysis. *Psychological bulletin*, *138*(2), 353-387.

Riglin, L., Petrides, K. V., Frederickson, N., & Rice, F. (2014). The relationship between emotional problems and subsequent school attainment: A meta-analysis. *Journal of Adolescence*, *37*(4), 335-346.

Ryan, R. M., & Deci, E. L. (2001). On happiness and human potentials: A review of research on hedonic and eudaimonic well-being. *Annual Review of Psychology*, *52*, 141-166. doi:10.1146/annurev.psych.52.1.141

APPENDIX A

Literature Search Strategy for the Broad Search of the Adolescent Mental Health and Well-Being

The following table presents the search terms used for the electronic database search:

ab,ti((universal NEAR/4 program*) OR (universal and intervention*) OR (universal and trial*) OR (universal and curricul*) OR (intervention NEAR/4 program*) OR (prevent* and program*) OR (prevent* and intervention*) OR (prevent* and trial*) OR (prevent* and curricul*) OR (school-based and program*) OR (school-based and intervention*) OR (school-based and trial*) OR (school-based and curricul*) OR ((skill* W/1 building) and program*) OR ((skill* W/1 building) and intervention*) OR ((skill* W/1 building) and trial*) OR ((skill* W/1 building) and curricul*) OR ("early intervention" OR "primary prevention" OR "secondary prevention" OR "prevention research" OR "group program*" OR "group intervention" OR "primary prevention" OR "secondary prevention" OR "tier 2" OR "tier two" OR "tier 1" OR "tier one" OR "multi-tier*" OR "selective prevention") OR (intervention NEAR/4 program) OR (prevention NEAR/4 research))
AND
ab,ti(school* OR student* OR "5th grade*" OR "fifth grade*" OR "grade* five" OR "grade* 5" OR "6th grade*" OR "Sixth grade*" OR "Grade 6" OR "Grade six" or "7th grade*" OR "Seventh grade*" OR "Grade* seven" OR "Grade* 7" or "8th grade*" OR "Eighth grade*" OR "Grade* 8" OR "Grade* eight" OR "9th grade*" OR "Ninth grade" OR "Grade* 9" OR "Grade* nine" OR "10th grade*" OR "Tenth grade*" OR "Grade* 10" OR "Grade* ten" OR "11th grade*" OR "Eleventh grade*" OR "Grade* 11" OR "Grade eleven*" OR "12th grade*" OR "Twelfth grade*" OR "Grade 12" OR "Grade* twelve" OR college* or universit* OR postsecondary OR post-secondary OR "institution* of higher learning" OR "two-year institut*" or "2-year institut*")
AND
ab,ti("control group*" OR random* OR "comparison group*" OR "matched group*" OR "treatment group*" OR experiment* OR evaluat* OR impact* OR effectiveness OR causal OR posttest OR post-test OR pretest OR pre-test OR QED OR RCT OR "propensity score" OR quasi-experimental OR efficacy OR "control condition*" OR "comparison condition*" OR "intervention group*" OR "intervention condition*" OR "no-intervention control" OR "wait-list" or "waitlist" or "waiting list" or waiting-list OR "intervention effect*" or "intervention children" OR "control children" OR postintervention or post-intervention OR preintervention OR pre-intervention OR "treatment class*" OR "intervention class*")

NOT
ab,ti(autis* OR kindergarten OR preschool OR “early reading” or “first grade” OR "pregnant women" OR "sickle cell" OR geriatric* OR dementia OR dental)

The databases searched were:

Academic Search Premier
EconLit with Full Text
Education Research Complete
ERIC
PsycINFO
Education Full Text (H.W. Wilson)
Social Sciences Full Text (H.W. Wilson)
Education Source
Dissertation Abstracts

Citations were selected from 1996 onward.

The search was conducted on January 5, 2016.