Supplementary Materials for: Dropout Prevention and Intervention Programs: Effects on School Completion and Dropout among School-aged Children and Youth
Sandra Jo Wilson, Emily E. Tanner-Smith, Mark W. Lipsey, Katarzyna Steinka-Fry, & Jan Morrison

Contents

DROPOUT PROJECT CODING MANUAL 2
REFERENCES FOR STUDIES INCLUDED IN THE META-ANALYSES 41
1.1 ELIGIBILITY CRITERIA

This meta-analysis deals with the effects of prevention and intervention programs on school completion/dropout. To be eligible for coding, a study must use an eligible intervention directed toward an eligible participant sample, report data that permit calculation of a numeric effect size for at least one eligible outcome variable, and employ an eligible research design.

1. INTERVENTIONS
   a. There must be a school-based or affiliated psychological, educational, or behavioral prevention/intervention program, broadly defined, that involves actions performed with the expectation that they will have beneficial effects on student recipients. School-based programs are those that are administered under the auspices of school authorities and delivered during school hours. School affiliated programs are those that are delivered with the collaboration of school authorities, possibly by other agents, e.g., community service providers, and which may take place before or after school hours and/or off the school grounds. Community-based programs that are explicitly presented as dropout prevention/intervention programs are eligible whether or not a school affiliation is evident. Other community-based programs that may include dropout among their goals or intended outcomes, but for which it or related variables are not the main focus, and which have no evident school affiliation are not eligible. Programs that are solely medical or pharmacological in nature are not eligible. Broad programs and policies that are at the district level where no intervention can be identified as occurring at the school level, such as district line restructuring, are not eligible.

2. SUBJECTS
   a. The research must investigate outcomes for an intervention directed toward school-aged youth, defined as those expected to attend pre-k to 12th grade primary and secondary schools, or the equivalent in countries with a different grade structure, corresponding to approximately ages 4-18. The age or school participation of the sample must be presented in sufficient detail to allow reasonable inference that it meets this requirement.
      • Recent dropouts who are between the ages of 18-22 are also eligible if the program under study is explicitly oriented toward secondary school completion or the equivalent.
   b. General population samples of school-age children are eligible. Samples from populations broadly at risk because of economic disadvantage, individual risk variables, and closely related factors are also eligible (e.g., inner city schools, students from low SES families, teen parents, students with poor attendance records, students who have low test scores or who are over-age for their grade). Samples consisting exclusively of specialized populations,
such as students with mental disabilities or other special needs, are not eligible. However, inclusion of some such individuals in a broader sample in which they are a minority proportion does not make that broader sample ineligible.

- Students with learning disabilities, such as dyslexia, that generally don’t require them to be in specialized schools or classrooms (i.e., they attend mainstream classes and typical schools) are considered eligible. NOTE: if studies with these types of samples are located, they should be set aside and brought to the attention of the group for discussion.

3. RESEARCH DESIGNS

a. An eligible study must use an experimental or quasi-experimental design; specifically, it must involve comparison of treatment and control conditions to which students are randomly assigned or non-randomly assigned with matching, statistical controls, or evidence of initial equivalence on key risk variables or student characteristics. The following research designs are eligible:

a. Participants were randomly assigned to treatment and control conditions or assigned by a procedure plausibly equivalent to randomization.

i. Participants in the treatment and control conditions were matched and the matching variables included a pretest for at least one qualifying outcome variable. However, if the qualifying outcome variable does not lend itself to meaningful pretest or the pretest values can be assumed zero, but the groups are matched on other variables plausibly related to risk for dropout, the study is still eligible. For this purpose, use of pretest or initial risk variables as statistical controls, e.g., in an ANCOVA or multiple regression analysis, is considered the equivalent of matching.

b. If participants were not randomly assigned or matched, the study must have both a pretest or relevant baseline risk variables and a posttest on at least one qualifying outcome variable with sufficient statistical information to derive an effect size or to estimate group equivalence from statements of statistical significance, or provide evidence of equivalence on key risk variables and/or student characteristics.

b. Studies that employ designs in which more than one treatment group is compared to a single control group are eligible; in these cases, effect sizes should be calculated for each treatment group compared to the control group.
Treatment–treatment studies that compare two or more treatments to each other without a control group may be eligible if one of treatment group receives a ‘sham’ or ‘straw-man’ treatment that is equivalent to a control condition, or if one of the treatments is a practice as usual condition in which that practice is not a distinctive program delivered at a relatively high level. E.g., if the school has a truancy officer engaged in routine activities for such a function, that would be acceptable as a practice as usual control condition. *NOTE: These should be set aside and brought to the attention of the group for discussion.*

Excluded designs:
Posttest only non-equivalent comparisons (not randomized, matched, or demonstrating equivalence) are not eligible. Single-group pretest-posttest designs are not eligible.

Each treatment and/or control group in a study must have at least 10 subjects at the time of assignment.

**4. OUTCOME VARIABLES**

a. The study must assess intervention effects on at least one outcome variable that represents school completion or dropout, or is a close proxy measure or recognized precursor for dropout. Qualifying outcome variables are those that fall in or are substantially similar to the following categories:
   - School completion/dropout;
   - GED completion/high school graduation;
   - Absences or truancy;
   - Enrollment/non-enrollment in school.

*NOTE: Studies in which the majority of children are under middle school age (approximately 5th grade or age 11) must have either a school completion or dropout outcome, or have attendance measures that are assessed in middle or high school.*

**5. DATE OF PUBLICATION**

- Eligible studies should be relatively modern, to be applicable to contemporary students. Therefore, the date of publication or reporting of the study must be 1985 or later even though the research itself may have been conducted prior to 1985. If, however, there is evidence in the report that the research was actually conducted prior to 1980 (more than five years before the 1985 cutoff date), then the study should be excluded.

**6. EFFECT SIZES**

- The study must report sufficient quantitative data to compute an effect size on an eligible outcome. In addition, the variables involved in the effect size must have a known direction of scoring, i.e., whether high or low values represent favorable or less favorable results. Studies that meet all eligibility criteria
except this, i.e., fall short only because an effect size cannot be calculated, should be identified and held separately for further consideration. NOTE: If a study meets all other eligibility criteria except for this one, do not exclude, but bring to the attention of the group for discussion.

7. STUDY SITE and LANGUAGE

- The study can be published in any language and conducted in any country as long as it meets all other eligibility criteria.
1.2 FILEMAKER & GENERAL CODING INFORMATION

There are six different FileMaker files that we use to code studies for this project. All of these files are linked together so that you can navigate between them by clicking the appropriate buttons. These files are defined below. Note: you may not use all of these files in your coding.

I. Bibliographic Database: We use this database to maintain the bibliography of potentially eligible reports. This database includes bibliographic information about each report, library location information, and tracking data about how far along each report is in the retrieval and coding process. You might use this database to search for related reports or to indicate when an article has been coded.

II. Eligibility Database: This database, generally accessed from the Bibliographic Database, contains information about the eligibility of each identified study; each study may be represented by one or more reports in the Bibliographic Database.

II. Header Database: This database includes the general information pertinent to a study. A study is defined as an investigation involving one independent group of subjects. In a quasi-experiment with a treatment and control group, the “study” includes both the treatment and control subjects. There will be one record in this database for each study that is coded (there may be multiple reports per study or multiple studies per report). Variables coded at this level include information about the methodology, treatment groups, initial group equivalence, etc.

III. Groups Database: This database includes information about the various groups that comprise a study. There will be one record in this database for each treatment AND comparison group for which there is sufficient data. Aggregate treatment and comparison groups are ALL groups in a study, including all levels of treatment and any type of comparison group. Thus, in a quasi-experimental study with one treatment group and one comparison group, you will have two records in this database. Variables coded at this level include information about the treatment and subjects under study.

IV. Dependent Variables Database: Within a given study, the researcher might evaluate the effectiveness of the intervention using multiple outcomes (or dependent variables). For example, the researcher might assess dropout during an observation period and the attendance days per subject. Thus, for this study, you would have two records in the dependent variables database, one for each outcome measure that you are coding. If the same outcome measure is used multiple times in a study, as in a pretest, a posttest, and a follow-up, you will only have one record for the outcome, but you will have multiple records for the effect sizes (see VI below). This database includes a description of each dependent variable, and some basic methodological information about the variable, such as reliability.
V. Breakouts: Breakouts are comparisons for subgroups of any treatment or control group, e.g., a treatment group compared with a comparison group using only the males in the sample. Each variable (e.g., gender, age) by which a group or groups are crossed constitutes one breakout; each value of that variable defines one subgroup; e.g., a male vs. female stratification is one breakout with two subgroups, one male and one female. If only the male subgroup is reported, there is still one breakout, but only one subgroup. Note that a simple report of the number of males and females in the treatment and control groups does not constitute a breakout (though it is relevant to group equivalence issues). To be a breakout, outcome data must be reported for the treatment-control comparison for at least one subgroup of the breakout variable. Breakouts are usually presented because the authors think that subgroups (e.g., males and females) are sufficiently different to warrant separate presentation of results (because, for example, males may be more likely to drop out than females).

VI. Effect Sizes: The effect size database tracks the actual statistical results of the study being coded. Because different researchers may present their results using different statistics (e.g., with a t-test or using means and standard deviations), we need to convert the statistics from each study into a common metric for our own analyses. The metric we use for this purpose is called an effect size. In most cases, the Effect Size database will calculate the appropriate effect size for the data that you enter. However, in some cases, you may use the Effect Size Determination Program (Excel Toolkit) or other resources to do the calculations. For each study you are coding, you will have one effect size record for EACH effect size – there can be anywhere from 1 to 50 or more effect sizes in one study and you will need to code each one separately. There will be more information below about how to do this coding.

You will also use the effect size database to record effect sizes related to the equivalence of treatment and control groups at the start of a study. As you know, it is vital that the two groups be as similar as possible at the beginning of the study on any characteristics, such as gender, age, and ethnicity, that might be related to dropout. When the two groups are similar, the differences between the two groups at the end of the study can be more easily attributed to the intervention and not these other characteristics. You will create a group equivalence record in the effect size database for each treatment-control comparison on a pretreatment variable that is relevant to group equivalence issues. That is, if the study provides you with the ages of the students in the treatment and comparison groups and the number of boys in each group, you will have two group equivalence records in the effect size database in addition to the regular effect sizes that you might code for this study.
1.3 HEADER VARIABLES CODING

**Step 1. Study Identifiers, Study Context, and Group Identification & Selection**

**STUDY IDENTIFIERS**

The “unit” you will code here consists of a study, i.e., one research investigation of a defined subject sample or subsamples compared to each other, and the treatments, measures, and statistical analyses applied to them. Sometimes there are several different reports (e.g., journal articles) about a single study. In such cases, the coding should be done from the full set of relevant reports, using whichever report is best for each item to be coded; **BE SURE YOU HAVE THE FULL SET OF RELEVANT REPORTS BEFORE BEGINNING TO CODE.** Sometimes a single report describes more than one study, e.g., one journal article could describe a series of similar studies done at different sites. In these cases, each study should be coded separately as if each had been described in a separate report.

Each study has its own study identification number, or StudyID (e.g., 619). Each report also has an identification number (e.g., 619.01), which you will find printed on the folder holding the report. The ReportID has two parts; the part before the decimal is the StudyID, and the part after the decimal is used to distinguish the reports within a study. (These two types of ID numbers, along with bibliographic information, are assigned and tracked using the bibliography.) When coding, use the study ID (e.g., 619) to refer to the study as a whole, and use the appropriate report ID (e.g., 619.01) when referring to an individual report.

While reading reports for coding, be alert to any references to other dropout studies that may be appropriate to include in this meta-analysis. If you find appropriate-looking references that are not currently entered into the bibliography, the references may need to be entered.

**[StudyID]** Study identification number of the study you are coding, e.g., 1923.

**[Coder]** Coder’s initials (select from menu)

**[H1]** Year of publication (four digits): If more than one report, choose earliest date.

**[CodeDate]** Date you began coding this study (will be inserted automatically)
STUDY CONTEXT

[H2] Country in which study conducted.
   1. USA
   2. Great Britain
   3. Canada
   4. Scandinavia: Denmark, Finland, Norway, Sweden
   5. Australia/New Zealand
   6. Other Western European Country: ___________
   7. Other: ____________________

[H3] Type of publication. If you are using more than one type of publication to code your study, choose the publication that supplies the effect sizes (in cases where more than one report provides effect sizes, choose a “peer reviewed” choice over another option, or choose the report that provides the most effect sizes).
   1. Book
   2. Journal article
   4. Thesis or dissertation
   5. Technical report
   6. Conference paper/presentation
   7. Other
   9. Cannot tell

GROUP IDENTIFICATION AND SELECTION

At this stage, you will need to identify the aggregate treatment and/or comparison groups used in the study for which effect size statistics can be computed. To do this, you will need to distinguish aggregate groups, which you will code here, from subgroups (or breakouts), which you will code later:

(1) Aggregate treatment and/or comparison groups. Aggregate treatment or control groups are the largest participant groupings on which contrasts between experimental conditions or contrasts between time points can be made. Note that the designations “comparison group” and “control group” refer to any group with which the treatment of interest is compared that is presumed to represent conditions in the absence of that treatment, whether a true random control or not. Often there is only one aggregate treatment group and one aggregate control group, but it is possible to have a design with numerous treatment variations (e.g., different levels) and control variations (e.g., placebos) all compared (e.g., in ANOVA format) to each other.
(2) Breakouts. Sometimes researchers will present data for some subset(s) of the participants from an aggregate group; e.g., for an aggregate group composed of males and females, the researchers may present some results for the males and females separately. You will code information about breakouts later.

Identifying the Aggregate Groups

Type in the name or identifier for each aggregate treatment group and each aggregate comparison group described in the study, whether you believe the group is eligible for coding or not.

Group labels used by researchers do not necessarily conform to the definitions of group types used in this project. In some cases, for example, researchers may compare one treatment with another treatment, and may call this “other” treatment a comparison or control group. For our purposes, if this “other” treatment group can realistically be expected to be effective, list it as a treatment group below; if it is a minimal or placebo treatment, not expected to produce an effect, list it as a comparison group.

Treatment Groups [H4a-d]
1 ___________________________
2 ___________________________
3 ___________________________
4 ___________________________

Comparison Groups [H5a-d]
1 ___________________________
2 ___________________________
3 ___________________________
4 ___________________________

[H4] Total number of treatment groups: ____

[H5] Total number of control groups: ____

ASSIGNMENT OF PARTICIPANTS

[H6] Unit of group assignment. The unit on which assignment to groups was based.
1. Individual (i.e., some children assigned to treatment group, some to comparison group)
2. Group (i.e., whole classrooms, schools, therapy groups, sites, residential facilities assigned to treatment and comparison groups)
3. Program area, regions, school districts, counties, etc. (i.e., region assigned as an intact unit)
9. Cannot tell
Method of group assignment. How participants/units were assigned to groups.

This item focuses on the initial method of assignment to groups, regardless of subsequent degradations due to attrition, refusal, etc. prior to treatment onset. These latter problems are coded elsewhere.

Random or near-random:
1. Randomly after matching, yoking, stratification, blocking, etc. The entire sample is matched or blocked first, then assigned to treatment and comparison groups within pairs or blocks. This does not refer to blocking after treatment for the data analysis.
2. Randomly without matching, etc. This also includes cases when every other person goes to the control group.
3. Regression discontinuity design: quantitative cutting point defines groups on some continuum (this is rare).
4. Wait list control or other quasi-random procedure presumed to produce comparable groups (no obvious differences). This applies to groups which have individuals apparently randomly assigned by some naturally occurring process, e.g., first person to walk in the door. The key here is that the procedure used to select groups doesn’t involve individual characteristics of persons so that the groups generated should be essentially equivalent.

Non-random, but matched: Matching refers to the process by which comparison groups are generated by identifying individuals or groups that are comparable to the treatment group using various characteristics of the treatment group. Matching can be done individually, e.g., by selecting a control subject for each intervention subject who is the same age, gender, and so forth, or on a group basis, e.g., by selecting comparison schools that have the same demographic makeup and academic profile of treatment schools.
5. Matched ONLY on pretest measures of some or all variables used later as outcome measures.
6. Matched on pretest measures AND other personal characteristics, such as demographics.
7. Matched ONLY on demographics: big sociological variables like age, sex, ethnicity, SES.

Nonrandom, no matching prior to treatment but descriptive data, etc. regarding the nature of the group differences:
8. Non-random, not matched, but pretreatment equivalence information is available.
99. Cannot tell
Confidence in assignment ratings. Overall confidence of judgment on how participants were assigned

1. Very low (little basis)
2. Low (best estimate)
3. Moderate (weak inference)
4. High (strong inference)
5. Very high (explicitly stated)

Equivalence of the groups being compared

At this point, you should go to the Effect Size Database to code group equivalence effect sizes and descriptive information about initial group differences for the study. See the Effect Size Coding Sheet section of this manual for more information on effect size calculation.

Number of variables on which treatment and comparison group differences were statistically compared prior to the intervention. A statistical comparison is one in which a statistical test was performed by the authors, whether they provide data or not (e.g., “no statistically significant differences were found”). Include in your count any demographic or risk factor comparisons as well as any comparisons on pretest variables, that is, measures of a dependent variable taken prior to treatment, e.g., prior number of absences when subsequent number of absences is used as an outcome measure.

Results of statistical comparisons.

1. No comparisons made
2. No statistically significant differences
3. Significant differences judged unimportant by coder. See note below regarding “importance” judgment.
4. Significant differences, judged of uncertain importance by coder
5. Significant differences, judged important by coder

Number of variables on which treatment and comparison group differences were or can be descriptively compared prior to the intervention. A descriptive comparison is any comparison across treatment and control groups that does not involve a statistical test (e.g., the actual number of males and females in each group or a statement by the author(s) about group similarity).

Results of descriptive comparisons.

1. No comparisons made or available
2. Negligible differences, judged unimportant by coder. See note below regarding “importance” judgment.
3. Some differences, judged of uncertain importance by coder
4. Some differences, judged important by coder
Note: An “important” difference means a difference on several variables relevant to the outcome variables, or on a major variable, or large differences; major variables are those likely to be related to dropout, e.g., SES, or family circumstances.

**[H13]** Rating of similarity of treatment and control groups. Using all the available information, rate the overall similarity of the treatment group and the comparison group, prior to treatment, on factors likely to have to do with dropout or responsiveness to treatment (ignore differences on any irrelevant factors). Note: Greatest equivalence from “clean randomization” with prior blocking on relevant characteristics and no subsequent attrition/degradation; least equivalence with some differential selection of one “type” of individual vs. another on some variable likely to be relevant to dropout.

Guidelines: Use ratings in the 1-3 range for good randomizations and matchings, e.g., 1=clean random, 2=nice matched. Use ratings in the 5-7 range for selection with no matching or randomization or instances where it has been seriously degraded, e.g., by attrition before posttest. Within this bracket, the question is whether the selection bias is pertinent to the outcomes being examined. Were participants selected explicitly or implicitly on a variable that might make a big difference in dropout? The middle three points are for sloppy matching designs, degradations, bad wait list designs, and the like. If the data indicate equivalence but the assignment procedure was not random give it a 4 or thereabouts since not all possible variables were measured for equivalence between groups.

1. Very similar, equivalent
2.
3.
4.
5.
6.
7. Very different, not equivalent

**[H14]** Overall confidence on rating of group similarity
1. Very low (little basis)
2. Low (best estimate)
3. Moderate (weak inference)
4. High (strong inference)
5. Very high (explicitly stated)

**[H15]** Click here to record any problems you encountered while coding this header.
1.4 GROUP EQUIVALENCE EFFECT SIZE CODING

At this point, you should go to the Effect Size Database to code group equivalence effect sizes and descriptive information about initial group differences. See the Effect Size Coding section of this manual for more information on effect size calculation.

For each measure you can identify on which the treatment and control group were compared prior to treatment (other than dependent variables) or on which you can tell equivalence (e.g. if all males then code it here), determine which group is favored and if possible, calculate an effect size (ES, standardized difference between means or odds ratio). Do not include here any comparisons on pretest variables, that is, measures of a dependent variable taken prior to treatment. In such cases the pretreatment ES is coded later as pretest information, not here as group equivalence information.

The only eligible variables for group equivalence effect sizes are: (a) gender, (b) age, (c) grade level, (d) race/ethnicity, and (e) variables relating to risk for school dropout. A pretest that is used later in the study as a posttest would not be coded here – you would code it as a pretest effect size. In matched group research designs, you will still code equivalence here for all eligible variables, even if groups were equally matched (e.g., both studies were 50% male, yielding a group equivalence effect size of .00). If the study reports group equivalence outcome data for multiple risk variables, group equivalence effect size information should be coded for up to four variables. If more than four variables are available for any of the risk factors, code the four most relevant ones. When deciding which are most relevant, use the following criteria:

1. First preference should be given to behavioral measures (e.g., prior absences, school performance).
2. Second preference should be given to measures of psychological conditions, predispositions, or attitudes (e.g., school engagement, school bonding, etc.).
3. Lowest preference should be given to broad measures of social disadvantage or family history (e.g., socioeconomic status of parents, residence in inner-city).

[StudyID] Indicate the Study ID for the study you are coding.

[ReportID] Enter the Report ID for the report in which you found the information on group equivalence. Use the complete Report ID, e.g. 1973.01.

[pagenum] Enter the page number on which you found the information on group equivalence.
Type of effect size:
5. Group Equivalence (for baseline treatment-control comparisons on variables other than the dependent variables)

Wave number. Pretests and group equivalence effect sizes always get a 1; each wave thereafter gets numbered consecutively, beginning with 1. Some studies involve more than one posttest measurement and we need to be able to distinguish one from another. Give the first posttest after treatment a 1, the second a 2, and so on.

Variable on which comparison is made:
____________________________ (e.g., gender, age, etc.)

Which group is favored? Whichever group has more of the characteristic that presumably makes them better off or more amenable to treatment (e.g., less truant, higher SES, smarter, etc.) is considered favored. NOTE: You should code this item even for cases in which you are unable to calculate a numeric effect size but have information about which group is favored.

1. Treatment (fewer males, younger, fewer minorities, less antisocial, less risk)
2. Control (fewer males, younger, fewer minorities, less antisocial, less risk)
3. Neither, exactly equal
9. Cannot tell, no report

Significance of group equivalence comparison (ONLY).
1. No statistically significant differences
2. Statistically significance differences
3. Negligible descriptive differences
4. Significant descriptive differences
98. N/A: No comparison made

Data Fields: Fill in the data fields using the relevant statistical information provided in the report(s). You do not need to fill in all the fields; fill in only the information necessary to calculate an effect size. Thus, if the report provides sample sizes, means, standard deviations, and t-test scores, you need only enter the sample sizes, means, and standard deviations.

ONCE YOU HAVE FINISHED CODING THE GROUP EQUIVALENCE EFFECT SIZE INFORMATION, YOU SHOULD RETURN TO THE HEADER FILE TO COMPLETE THE CODING OF HEADER VARIABLES.
1.5 TREATMENT AND CONTROL GROUPS CODING

Create one record in this database for each of the aggregate treatment and/or control groups that you selected earlier for coding. Studies with a treatment group and a control group will have two records, etc.

Group Identification and General Nature Of Treatment

[StudyID] Type in the StudyID for the study you are coding if it does not appear automatically.

[GroupID] Number each group consecutively within a study, starting with 1.

[G1] Select the type of group you are coding.
   1. Treatment group
   2. Control group

[G2] What general type of “treatment” does this group receive?

Intervention Condition

1. Focal program or treatment. There may be several focal programs in a study, as when two different types of treatments, both of which could be expected to be effective, are compared.

Control Condition

2. “Straw man” alternate program or treatment, diluted version, less extensive program, etc., not expected to be effective but used as contrast for treatment group of primary interest. If the alternate treatment is not minimal and could realistically be expected to be effective, it is not a control condition and should be classified as a focal treatment instead.

3. Placebo (or attention) treatment. Group gets some attention or sham treatment (e.g., watching Wild Kingdom videos while treatment group gets therapy)


[G3] Program name. Write in program or treatment label for this group (e.g., Dropout Prevention Curriculum, waiting list control, etc.). REMEMBER: YOU MUST CREATE A PROGRAM LABEL FOR CONTROL GROUPS AS WELL AS TREATMENT GROUPS.

[G4] Program description. Write in a brief description of the treatment this group receives. Please try to keep the description short by focusing on the key elements of
treatment, but make sure you include ALL treatment elements in your description. As much as possible, quote or give a close paraphrase of the relevant descriptive text in the study report. REMEMBER: YOU MUST CREATE A DESCRIPTION FOR CONTROL GROUPS AS WELL AS TREATMENT GROUPS.

**TREATMENT CHARACTERISTICS**

[G5] Intervention type:
1. School-based (administered under the auspices of school authorities and delivered during school hours)
2. School affiliated (delivered with the collaboration of school authorities, possibly by other agents, e.g., community service providers; may take place before or after school hours and/or off the school grounds)
3. Community-based (explicitly presented as dropout prevention/intervention programs; may or may not have a school affiliation)
4. Not applicable (control condition)


*For each treatment AND control condition:*
First check all program types that apply to a given intervention (e.g., a program may include GED preparation, cognitive behavioral techniques, tutoring, and contingency management).

Second, choose the **one** program type that can be considered the focal program characteristic. Most programs will arguably deliver multiple service types, but do your best to narrow the focal type down to one category. It may be helpful to examine the amount of each service type delivered. For instance, if a program delivered 1 hour/week of skills training to parents and 5 hours/week of vocational training to students, you would code vocational training as the focal program component. If a program contains too many service types to distinguish a focal type, choose “multi-service” package as the focal component.
ACADEMIC:
1. Curriculum
2. ESL/ELL (English as a second language/ English language learners)
3. Remedial education (e.g., reading remediation)
4. GED preparation
5. Computer-assisted learning
6. Test-taking and study skills assistance
7. Tutoring
8. Homework assistance
9. Extracurricular activities (e.g., after school club). NOTE: just because a program is delivered after school does not mean it should be coded here; this program component should include academic, social, or sport activities that are separate from regular school activities.
10. Professional development for school staff
49. Individualized teaching

SCHOOL STRUCTURE
11. Class or grade reorganization (schools within schools, team teaching)
12. Small class sizes/small “learning communities”
13. Alternative school (e.g., small school settings comprised primarily of students with severe academic or behavioral problems that preclude them from attending regular classes; i.e., this is the ‘last chance’ for many students who may have otherwise been expelled or suspended from school)

FAMILY ENGAGEMENT:
14. Family outreach
15. Feedback to parents and students on performance
16. Parent or teacher consultation enhancement
17. Parenting skills program
47. Skills training for significant others

COLLEGE FOCUSED/CONNECTING STUDENTS TO ATTAINABLE FUTURE:
18. Academic advising
19. College-preparatory curriculum
20. Academic summer/weekend program (i.e., enrichment programs)
21. College campus visits
22. College and financial aid application assistance
23. College scholarships

WORK RELATED/ FINANCIAL SUPPORT:
24. Internships
25. Career exploration
26. Vocational training
27. Job placement assistance
28. Living allowance
29. Bonuses and sanctions applied to welfare grant

LINKING TO SERVICES:
30. Case management
31. Health services
32. Transportation assistance
33. Child care/day care
34. Residential living services

SOCIAL RELATIONSHIPS:
35. Mentoring
36. Peer support
37. Social events
38. Community service/volunteer service/student as tutor (“helper-therapy”)  
39. Recreational, wilderness, etc. program

PERSONAL/AFFECTIVE:
40. Counseling
41. Skills training (life skills, social skills/social competence)
42. Cognitive behavioral therapy (e.g., problem solving skills)

BEHAVIORAL:
43. Attendance monitoring
44. Contingency management, financial incentives, token economy, extrinsic reward system to promote attendance/academic achievement

OTHER:
45. Multi-service package (NOTE: Only choose this program code if the group receives an amorphous, broadly defined program with components that cannot be clearly identified otherwise. Use this program code as focal if a group has multiple “focal” treatment components and you cannot make a distinction otherwise.

46. OTHER (Please, describe [prog50a]______________)  

88. control group
**[G9] Treatment Site.** Nature of the site in which treatment generally delivered: (select one)

**School Sites**
1. Regular Class Time (this includes interventions delivered during regularly scheduled classes AND in the regular classroom for youths in the group)
2. Special Class (e.g., youth in treatment are in a classroom-type setting that is different from a typical classroom, although it may be the subjects’ usual classroom – includes such settings as special education classrooms, schools-within-schools, alternative schools, etc.)
3. Resource Room, School Counselor’s Office, or other similar setting that is NOT the student’s regular classroom; the idea here is that students are removed from class for treatment
4. Treatment delivered at school facility, but not during regular school hours (e.g., afterschool programs)

**Home**
5. Treatment delivered in the subject’s home

**Community-based, Non-residential**
6. Private office, clinic, center (e.g., YMCA, university, therapist’s office)
7. Public office, clinic, center (e.g., human services department, public health agency)
8. Work site (e.g., community service, trash collection on roadside, etc.)
9. Park, playground, wilderness area, etc.

**Institutional, Residential**
10. Private institution, residential
11. Public institution, residential (e.g., camp, reformatory)

**Mixed or Multiple Sites**
12. School and home
13. Other mixed, some combination of above sites

88. N/A: control group
99. Cannot tell

**[G10] Role of the evaluator(s)/author(s)/research team or staff in the program.**
This item focuses on the role of the research team working on the evaluation, regardless of whether they are all listed as authors.
1. Evaluator delivered therapy/treatment
2. Evaluator involved in planning or controlling treatment or is designer of program
3. Evaluator influential in service setting but no direct role in delivering, controlling, or supervision
4. Evaluator independent of service setting and treatment; research role only
8. Not applicable, control condition
9. Cannot tell

[G11] Role of program developer in the research project. This item focuses on the individual (or group of individuals) who created or developed the program and their role in the delivery of the program under study. Is the program developer the researcher conducting the study, or is the program developer not participating in the research project?
1. Program developer is author/evaluator/delivery agent
2. Delivery agent/author/evaluator modified existing program, but original program developer is not involved (note: this response suggests that the author/evaluator/delivery agent takes on a sort of quasi-developer status by modifying a program)
3. Program developer is not affiliated with research study and program is delivered as originally intended by developer
8. Not applicable, control condition
9. Cannot tell

[G12] Routine practice or program vs. research project. Indicate the appropriate level for the treatment you are coding: at one end of the continuum are research projects (option 1), in which a researcher decides to implement and evaluate a particular program for research purposes; in many cases, the program may require the cooperation of a service agency (school, clinic, etc.), but the intervention is delivered primarily so the researcher can conduct research. At the other end of the continuum are evaluations of “real-world” or routine programs (option 3): a service agency implements a program on its own, and also decides to conduct an evaluation of the program; the evaluation may or may not be conducted by outside researchers. In the middle of the continuum are demonstration projects (option 2), which are conducted primarily for research purposes, but generally have more elements of “real world” practice than typical research projects as defined under option 1. Demonstration projects generally involve a program that has been studied in prior research but is being tested for effectiveness in different settings than the original research, or on a larger scale than the original research.

If a researcher is a school principal and is conducting the evaluation as part of his/her dissertation, the decision depends on the extent of the program. If the program is small-scale and implemented in, say, a classroom or two, and supervised by the researcher/principal, code it as a research project. If the program is a broader school-wide program that the researcher/principal happens to be evaluating, code it as either a demonstration or routine program, depending on whether the program is
a special program being tested (demonstration) or something that the school does on a routine basis (routine practice).

1. **Research project**: The intervention would not have been implemented without the interest or initiative of the researcher(s). The intervention is delivered by the research staff or by service providers (regular agency personnel, teachers, etc.) trained by the researchers.

2. **Demonstration project**: A research project that involves a new or special program being tested, rather than a routine program. Although generally implemented by researchers for research purposes, a demonstration project has more elements of actual practice than a research project. Demonstration projects usually involve programs that have been studied previously, either in small-scale pilot projects or tightly controlled efficacy trials; demonstration projects would serve as a larger scale or quasi-real-world test of a promising program.

3. **Evaluation of a “real-world” or routine program**: A service agency implemented the program using routine personnel and the typical clients for that program; there may be outside researchers who conduct the evaluation, but the program they are evaluating was already in place before the research began and is presumed to continue after the research has ended.

8. Not applicable, control condition
9. Cannot tell

**[G13]** Treatment provider’s discipline. Indicate the discipline or type of treatment provider for the treatment. This item focuses on the individual(s) who have direct contact with the subjects in treatment, not necessarily the persons conducting the data analysis or evaluation. In multi-service treatment programs with multiple providers, indicate the discipline of the individual(s) who provide the focal or modal treatment modality.

1. Teacher
2. School guidance counselor
3. School psychologist
4. School personnel, other than school counselor or teacher (e.g., principal, school nurse)
5. Counselor
6. Social worker
7. Researcher or researcher’s staff, graduate students
8. Other
88. N/A: no treatment received
99. Cannot tell
Did treatment personnel receive special training in this specific program, intervention, or therapy? If the treatment is delivered by the researcher, use “yes” below, unless the report indicates otherwise.

1. Yes
2. No
9. Cannot tell

If yes, write in amount of training of personnel for providing this treatment:
________________________

Treatment Format:

For each treatment AND control condition:
First check all formats that apply to a given intervention (e.g., a program may include group and individual components, or have a family component).

Second, choose the one format type that can be considered the focal format. This selection should match the format of the focal program type you selected above under G6. If you selected multi-service package above, select the format for the most frequent or most focal piece of the package; if this is impossible, select multiple format program.

[1] ___ Subject alone (self-administered treatment)
[2] ___ Subject & provider, one-on-one
[3] ___ Subject group and provider, not classroom
[4] ___ Subject group and provider, classroom
[5] ___ Parents only and provider, child not present
[6] ___ Group of parents and provider, children not present
[7] ___ Child & parents with provider
[8] ___ Group of families with provider
[9] ___ Child & parents, no provider (self-administered treatment)
[10] ___ Teachers, treatment professional, no children
[12] ___ Multiple format program; no focal format
[88] ___ N/A: control group

Focal Treatment Implementation/Length/Integrity

Duration of treatment. Approximate (or exact) number of weeks that subjects received treatment, from first treatment event to last excluding follow-ups designated as such. Divide days by 7; multiply months by 4.3. Code 777 if a control group that receives nothing Code 999 if cannot tell. Estimate for this item if necessary, and if you can come up with a reasonable order of magnitude number. Use school year conversions listed below.
[G22] Approximate (or exact) frequency of contact between subjects and provider or treatment activity. This refers only to the element of treatment that is different from what the control group receives.

1. Less than weekly
2. Once a week
3. 2 times a week
4. 3-4 times a week
5. Daily contact (not 24 hours of contact per day but some treatment during each day, perhaps excluding weekends)
6. Continuous (e.g. residential living)
9. Cannot tell
88. N/A: control group

[G24] ______________ Approximate (or exact) mean hours actual contact time between subject and provider or treatment activity per week if reported or calculable. Assume that high school classes, counseling, or therapy sessions are an hour unless otherwise specified. Round to one decimal place. Code 7777 for control groups that receive nothing; 8888 for institutional, residential, or around the clock program; code 9999 if not available. Use school year conversions listed below.

[G26] ______________ Approximate (or exact) mean number of hours total contact between subject and provider or treatment activity over full duration of treatment per subject if reported or calculable. Round to whole number. Code 7777 for control groups that receive nothing; 8888 for institutional, residential, or around the clock program; code 9999 if not available. Use school year conversions listed below.

[G28] Were there additional untimed treatment components that were not included in the dosage estimates given above? For example, these could be wrap-around or other diffuse services like case management that aren’t presented in enough detail in the study reports to estimate dosage.

1. Yes
2. No

[G51] Were the dosage estimates given above for a treatment program that was delivered to significant others of the target subjects, rather than the subjects themselves?

1. Yes
2. No

School Year Conversions
Length of School Year (i.e., duration) (approx):
1 school year = 4 quarters = 2 semesters = 9 months = 38.7 weeks = 271 days
Actual TIME IN CLASS (i.e., contact) (approx.):
1 school year = 8.4 months = 36 weeks = 180 days
1 semester = 4.2 months = 18 weeks = 90 days
1 quarter = 2 months = 9 weeks = 45 days

Misc. Conversions
Hours in a school day for STUDENTS: 6.5 - 7 hours
Hours in a school day for TEACHERS: 8 hours
Typical class period (High School): 45 min- 1 hour
Typical therapy session (counseling): 1 hour

<table>
<thead>
<tr>
<th>Instructional Days per School Year</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>U.S. public schools average</td>
<td>180 days</td>
</tr>
<tr>
<td>States with less than 180</td>
<td>Arkansas, Colorado, Illinois, Kentucky, Louisiana, Maine, Missouri, North Dakota, Oklahoma, Vermont, Wyoming (average 174 days)</td>
</tr>
<tr>
<td>States with more than 180</td>
<td>Kansas (K-11, 186 days; Grade 12, 181 days) and Ohio (182 days)</td>
</tr>
<tr>
<td>U.S. private schools average</td>
<td>180.4 days</td>
</tr>
<tr>
<td>International average</td>
<td>193 days</td>
</tr>
<tr>
<td>Korean average</td>
<td>225 days</td>
</tr>
<tr>
<td>Japanese average</td>
<td>223 days</td>
</tr>
<tr>
<td>Chinese average</td>
<td>221 days</td>
</tr>
<tr>
<td>Australian average</td>
<td>196 days</td>
</tr>
<tr>
<td>Russian Federation average</td>
<td>195 days</td>
</tr>
<tr>
<td>Netherlands average</td>
<td>191 days</td>
</tr>
<tr>
<td>English average</td>
<td>190 days</td>
</tr>
<tr>
<td>Canadian average</td>
<td>188 days</td>
</tr>
</tbody>
</table>

Instructional Hours per School Year
(*Eight U.S. states do not set a minimum number of instructional days; instead they set number of instructional hours.*)

<table>
<thead>
<tr>
<th>Delaware</th>
<th>Grades 1-11 (1060 hours); Grade 12 (1032 hours)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Idaho</td>
<td>Grades 1-3 (810 hours); Grades 4-8 (900 hours); Grades 9-12 (990 hours, includes 22 hours for staff development)</td>
</tr>
<tr>
<td>Michigan</td>
<td>1080 hours</td>
</tr>
<tr>
<td>Montana</td>
<td>Grades K-3 (720 hours); Grades 4-12 (1080 hours, 1050 for graduating seniors)</td>
</tr>
<tr>
<td>State</td>
<td>Grades</td>
</tr>
<tr>
<td>----------------------</td>
<td>-------------------------------------------</td>
</tr>
<tr>
<td>Nebraska</td>
<td>Grades 1-8 (1032 hours); Grades 9-12 (1080 hours)</td>
</tr>
<tr>
<td>Oregon</td>
<td>Grades 1-3 (810 hours); Grades 4-8 (900 hours); Grades 9-12 (990 hours, seniors hours may be up to 30 less)</td>
</tr>
<tr>
<td>South Dakota</td>
<td>Grades 4-12 (962.5 hours)</td>
</tr>
<tr>
<td>Virgin Islands</td>
<td></td>
</tr>
<tr>
<td>Average – all grades</td>
<td></td>
</tr>
<tr>
<td>Average – high schools</td>
<td></td>
</tr>
<tr>
<td>Average of states with day and hour minimums</td>
<td></td>
</tr>
<tr>
<td>All grades</td>
<td></td>
</tr>
<tr>
<td>High schools</td>
<td></td>
</tr>
</tbody>
</table>

- If the intervention is only one class (math, study skills, etc.) just count one hour per day for the dosage.
- If the whole school experience in the intervention (alternative school, smaller classes, etc.) use the charts above to calculate dosage.

**[G29] Monitored treatment implementation.** Was the implementation of the program monitored by the author/researcher or program personnel to assess whether it was delivered as intended?

1. Yes. Do not infer that monitoring happened. Select “yes” ONLY if the report specifically indicates that implementation was monitored.

2. Possible
d
3. No, apparently implemented as intended

**[G30] Based on evidence or author acknowledgment, was there any uncontrolled variation or degradation in implementation or delivery of treatment, e.g., high dropouts, erratic attendance, treatment not delivered as intended, wide differences between settings or individual providers, etc.? Assume that there is no problem if one is not specified.**

This question has to do with variation in treatment delivery, not research contact. That is, there is no “dropout” if all subjects complete treatment, even if some fail to complete the outcome measures.

1. Yes (describe below)
2. Possible (describe below)
3. No, apparently implemented as intended

**[G31] Implementation Monitoring Procedures and Problems.** Describe any implementation problems or issues mentioned by the authors. Also describe any procedures used to monitor implementation fidelity.
Subject Characteristics

[G40] Gender composition of group.
1. No males (<5%)
2. Some males (<50%)
3. 50% to 60% male
4. Mostly males (>60%)
5. All males (>95%)
9. Cannot tell

[G42] Enter percent male: ___________ (use decimal rather than whole number, i.e., .42 NOT 42%)

ETHNICITY CODING (Code 9999 if you cannot tell)
[G43a] Percent white.
[G43b] Percent black
[G43c] Percent Hispanic
[G43d] Percent other minority
[G43e] Percent non-white (ONLY use this category if specific minority groups are not mentioned; if you use this category, there should only be numbers in the white and non-white categories)

Rankings: 1=clear majority; 2=present but proportion unknown; 3=clear minority; 0=not present.

[G44a] White rank
[G44b] Black rank
[G44c] Hispanic rank
[G44d] Other minority rank
[G44e] Non-white rank (ONLY use this category if specific minority groups not mentioned; if you use this category, there should only be numbers in the white and non-white categories)

[G45] Describe others and/or non-whites:
______________________________________.

[G46] Enter the average age of the sample using number of years. Enter 9999 if you cannot tell.

[G46a-b] High and low age using years. Enter 9999 if you cannot tell.

[G47] Enter the average grade level of the sample. (dropdown menu)
[G47a-b] High and low grades (dropdown menu)

[G48] Predominant level of “risk” of youths in the sample:

Think of the reason that the subjects in this group ended up in this group; did the researchers select potential dropouts for treatment; if yes, how were the potential dropouts identified?

[G49] Socioeconomic status: Type in a brief description of the socioeconomic composition of the sample. This might include information on the percentage of children eligible for free lunches, the income level of the children’s parents, or a description of poverty in the community. Quote or closely paraphrase the relevant descriptive information in the report.

[G50] Please describe any problems you encountered while coding this record.

1.6 DEPENDENT VARIABLES CODING

Select the general construct group for the dependent variable you are coding, then select the specific construct category that best matches the dependent variable.


[DV1] Construct Group

100. Dropout
101. Attendance, truancy
102. Academic performance
103. School conduct
104. School engagement

[DV2] Specific Construct

Dropout

200. Dropout
201. Graduation
202. GED completion
203. Enrolled in post-secondary education
221. Graduated OR obtained GED

Attendance

204. Absences/truancy
205. Tardies
206. Attendance
222. Enrolled in high school; attending or not attending
**Academic performance**
207. GPA, grades
208. Standardized test scores
209. Academic track
210. Grade retention
211. Unstandardized, generic academic achievement score
220. Academic credits (# earned, average credits)

**School conduct**
212. Suspensions
213. Expulsions
214. Detention
215. Classroom behavior

**School engagement**
216. School self-concept
217. Academic expectations/goal setting
218. Attitude toward school/school bonding
219. Attitude toward teachers

**Employment**
223. Any employment (part-time, full-time)
224. Full-time employment
225. Hours worked per week

**[DV3]** Source of information. Who provided the information for this dependent variable?
1. Participants, self-report
2. Parents
3. Peers
4. Teachers
5. Principal
6. Therapist/Service Provider (treatment agent)
7. School Records
8. Researcher or interviewer
9. Involved other (not treatment agent, not researcher), e.g., school counselor.
10. Multiple sources, cannot tell which is dominant
99. Cannot tell

**[DV4]** Type of Measure.
1. Survey, questionnaire, or interview
2. Standardized test (e.g., standardized achievement test)
3. School records
4. Other: __________
9. Cannot Tell

[DV6] Time period covered by this dependent variable.

Total number of weeks over which the information presented in this dependent variable was counted. This question applies mainly to variables like attendance that are continuously counted and thus might be presented in study reports as: number of absences in the past month (you would code 4.3 weeks for this) or weekly attendance over the past semester (you would code the number of weeks in the semester). Measures like dropout or graduation, which are measured at discrete time points and do not cover a certain time period, should be coded as 888 for not applicable.

If you have two measures of the same construct (such as attendance) that have different time period coverage, then you must create two separate dependent variables.
1.7 BREAKOUT/SUBGROUP CODING

Breakouts are comparisons involving subgroups of an aggregate treatment and/or control group. For example, the males in a treatment group might be compared with the males in a comparison group, or pretest-posttest results might be presented for males and females separately. Each variable (e.g., gender, age) by which the aggregate group(s) are subdivided constitutes one breakout, and each value of that variable defines one subgroup; i.e., a males vs. females stratification is one breakout (gender) with two subgroups, one male and one female. If only the male subgroup is reported, there is still one breakout, but only one subgroup.

Note that a simple report of the number of males and females in the treatment and control groups does not constitute a breakout (though it is relevant to group equivalence issues). To be a breakout, outcome data must be reported for the treatment-control or pretest-posttest comparison for at least one subgroup of the breakout variable. Breakouts are usually presented because the authors think that subgroups (e.g., males and females) are sufficiently different to warrant separate presentation of results (because, for example, males may exhibit more aggressive behaviors than females).

NOTE: Only certain breakout variables are eligible for coding. These include gender, age, ethnicity, and prior school completion/dropout, GED completion, or absences/truancy. If you encounter another breakout variable that may be relevant to dropout, please check with Sandra.

Create a new record for each subgroup that you will be coding for this study.

[StudyID] Study ID for the study you are coding.

[BreakID] Subgroup number. Assign a number to the subgroup such that the first subgroup you code is numbered 1, the second is numbered 2, and so on. These numbers are used within a study, so when you code subgroups from another study, you would start over with 1 again.

[Labels:B2] Write in descriptor for the subgroup you are coding, e.g., males, 8 year olds, whites, etc.
1.8 EFFECT SIZE CODING

Although this is the final section of coding, it is a good idea to identify at least one codable effect size before you start coding a study, because studies that appear eligible frequently end up presenting data that cannot be coded into an effect size.

This portion of coding requires familiarity with some basic statistics, including means, standard deviations, proportions, t-tests, chi-squares, ANOVA (or F-tests), and the like.

Step 1. General Information

[StudyID] Type in the appropriate StudyID if it does not appear automatically.

[ReportID] Report ID for this effect size. Indicate the report number (e.g., 2098.01) for the report in which you found the information for this effect size. This is important so that we can find the source information for the effect sizes later on, if necessary, and is especially important for studies with multiple reports.

[ESID] Effect size ID. FileMaker will automatically generate unique effect size ID numbers ACROSS studies.

[pagenum] Page number for this effect size. Indicate the page number of the report identified above on which you found the effect size data. If you used data from two different pages, you can type in both, but use a comma or dash between the page numbers.

There are 3 types of effect sizes that can be coded: pretest, posttest, and group equivalence (or baseline similarity) effect sizes. They are defined as follows:

- **Pretest effect size.** This effect size measures the difference between a treatment and comparison group before treatment (or at the beginning of treatment) on the same variable used as an outcome measure, e.g., school attendance measured before the treatment begins is used as a pretest for school attendance measured the same way after the treatment ends.

- **Group equivalence effect size.** Group equivalence effect sizes are used to code the equivalence of two groups prior to treatment delivery on variables that might be related to outcome. See the Group Equivalence Coding section for more information.

- **Posttest effect size.** This effect size measures the difference between two groups after treatment on some outcome variable.
This is very important!!!! These three types of effect sizes are different from the multiple breakouts and multiple dependent variables that you might have in a study. For example, you might have a study that measures the treatment and comparison groups at pretest and posttest at 6 months after treatment on 3 different dependent variables. The results might be presented for the entire sample and broken down by gender. In this case you would have 6 group comparison effect sizes for the entire sample – three for the pretest and 3 for the 6 month posttest (the three is for your three dependent variables). In addition to these 6 aggregate effect sizes, you will have 6 more for the girls (the same as for the aggregate groups but just for the subgroup of girls) and 6 for the boys (also the same as for the aggregate groups but just for the subgroup of boys).

[ES24] Type of effect size:
1. Pretest (for treatment-control baseline comparison on a dependent variable)
2. Posttest (for the first treatment-control outcome comparison on a dependent variable)
5. Group Equivalence (for baseline treatment-control comparisons on variables other than the dependent variables)

[ES19] Wave number. Pretests and group equivalence effect sizes always get a 1; each wave thereafter gets numbered consecutively, beginning with 1. Some studies involve more than one posttest measurement on the same dependent variable, and we need to be able to distinguish one from another. Give the first posttest after treatment a 1, the second a 2, and so on.

[ES47] Timing of measurement. Approximate (or exact) number of weeks after treatment when measure was taken. Divide days by 7; multiply months by 4.3. Enter 999 if cannot tell, but try to make an estimate if possible. Enter 0 if pretest. If posttest measurement occurred during an ongoing treatment, use 888 here.

[es47_ck]

Step 2. Group Selection

[GroupID1] Group 1
If you are coding a treatment-control effect size, select the appropriate treatment group here. If you are coding a treatment-treatment effect size, select the focal treatment group here or, if neither is focal, select one here and the other as Group 2 below.

[GroupID2] Group 2
If you are coding a treatment-control effect size, select the appropriate control group here. If you are coding a treatment-treatment effect size, select the second of the two treatment groups here.

[BreakID] Select Breakout group if relevant.
Step 3. Dependent Variable Selection

[VarNo] Select the dependent variable for this effect size.

Step 4. Effect Size Calculation and Data Entry

It is now time to identify the data you will use to calculate the effect size and to calculate the effect size yourself if necessary (see below). Effect sizes can be calculated ONLY from data based on the number of subjects, e.g., average number of days absent per subject and the corresponding standard deviation) or proportion of subjects who were chronic truants during a given time period. Effect sizes can NOT be calculated from data based solely on the incidence of events, e.g., total number of days absent per group. THIS IS VERY IMPORTANT—BE SURE YOU KNOW WHICH KIND OF DATA YOU HAVE.

You need to determine what effect size format you will use for each effect size calculation. There are two general formats you can use, each with its own section in FileMaker:

1. Compute ES from means, sds, variances, test statistics, etc.
2. Compute ES from frequencies, proportions, contingency tables, odds, odds ratios, etc.

Also note that within each of the above effect size formats, effect sizes can be calculated from a variety of statistical estimates; to determine which data you should use for effect size calculation, please refer to the following guidelines in order of preference:

1. Compute ES from descriptive statistics if possible (means, sds, frequencies, proportions).
2. If adequate descriptive statistics are unavailable, compute ES from significant test statistics if possible (values of t, F, Chi square, etc.).
3. If significance tests statistics are unavailable or unusable but p value and degrees of freedom (df) are available, determine the corresponding value of the test statistic (e.g., t, chi-square) and compute ES as if that value had been reported.

Note that if the authors present both covariate adjusted and unadjusted means, you should use the covariate adjusted ones. If adjusted standard deviations are presented, however, they should not be used.
Which group is favored?

Select the group that has done “better”:
1. Treatment
2. Control
3. Neither, exactly equal
4. Cannot tell

For treatment-control comparisons, the treatment group is favored when it does “better” than the control group. The control group is favored when it does “better” than the treatment group.

Remember that you cannot rely on simple numerical values to determine which group is better off. For example, a researcher might assess the attendance and report this variable in terms of the average number of absences in the last semester. Fewer absences are better than more, so in this case a lower number, rather than a higher one, indicates a more favorable outcome.

Sometimes it may be difficult to tell which group is better off because a study uses multi-item measures in which it is unclear whether a high score or a low score is more favorable. In these situations, a thorough reading of the text from the results and discussion sections usually can bring to light the direction of effect – e.g., the authors will often state verbally which group did better on the measure you are coding, even when it is not clear in the data table. Note that if you cannot determine which group has done better, you will not be able to calculate a numeric effect size. (You will still be able to create an effect size record—just not a numeric effect size.)

Effect size derived from what type of statistics?
1. Means and SDs; means and variances; means and standard errors
2. N successful/unsuccessful (frequencies)
3. Proportion successful/unsuccessful (percentage successful or not)
4. Multi-category (polychotomous) frequency or %
5. Independent t-test
6. One-way ANOVA (2 groups, 1 degree of freedom)
7. One-way ANOVA (>2 groups, >1 degree of freedom)
8. Covariance Adjusted (ANCOVA)
9. Chi-square statistic (1 degree of freedom; from 2x2 table)
10. Correlation coefficient (zero-order)
11. Hand calculated ES
17. Effect sizes as reported directly in the study
18. Other (please specify)
For this effect size, did you use adjusted data (e.g., covariate adjusted means) or unadjusted data? If both unadjusted and adjusted data are presented, you should use the adjusted data for the group means or mean difference, but use unadjusted standard deviations or variances. Adjusted data are most frequently presented as part of an analysis of covariance (ANCOVA). The covariate is often either the pretest or some personal characteristic such as socioeconomic status. If you encounter data that is adjusted using something other than a covariate, please see Sandra or Mark.

1. Unadjusted data
2. Pretest adjusted data (or other baseline measure of an outcome variable construct)
3. Data adjusted on some variable other than the pretest (e.g., socioeconomic status)
4. Data adjusted on pretest plus some other variables

Confidence in effect size calculation

1. High estimate (e.g., have N and crude p values only, e.g., p<.10, and must reconstruct via rough t-test equivalence)
2. Moderately estimated (e.g., have complex but relatively complete statistics, e.g., multiple regression, LISREL, multifactor ANOVA, etc. as basis for estimation)
3. Some estimation (e.g., have unconventional statistics and must convert to equivalent t-values or have conventional statistics but incomplete, such as exact p values only)
4. Slight estimation (e.g., must use significance testing statistics rather than descriptive statistics, but have complete statistics of the conventional sort, such as a t-value or F-value)
5. No estimate (e.g., have descriptive data: means, sds, frequencies, proportions, etc.; can calculate an ES directly.)

Significance information for this comparison.
For treatment-control and treatment-treatment comparisons: Did the authors provide any information about the statistical significance of the difference between the two groups you selected on the dependent variable you selected for the time point you have selected for this comparison? Sometimes authors will state that a particular comparison was not significant, but not provide any calculable effect size data. In these cases, you should select “5” for this item. The effect size field should remain blank. In other cases, authors will state that a particular comparison was significant, but not provide any calculable effect size data. In these cases, you should select “4” for this item. Again, the effect size field should remain blank. NOTE: the last three options (4, 5, and 6) are for cases for which you have direction (i.e., you know which group is favored) but no effect size information.
1. Significant result, ES data below
2. Non-significant result, ES data below
3. Significance not reported, ES data below
4. Significant result, no ES data
5. Non-significant result, no ES data
6. Significance not reported, no ES data

[ES55] Intent-to-treat analysis: Are results for this effect size based on an intent-to-treat analysis?

Experimental and quasi-experimental designs may employ “intent-to-treat” (ITT) or “completer” analyses. An intent-to-treat analysis is one that (attempts to) includes all randomized subjects in the groups to which they were randomly assigned, regardless of the compliance with the entry criteria, the treatment they actually received, or any subsequent withdrawal from treatment or deviation from the protocol. A true ITT is possible only when the authors (attempt to) use outcome data for all randomized subjects; if all assigned subjects are used to present outcome results, then code as ITT, regardless of whether authors call the analysis an ITT. If the authors attempt to collect outcome data on non-completers and even if they are not 100% successful in this attempt, still code as ITT (as the missing data for non-completers is due to attrition). Sometimes researchers will use a modified ITT, in which they estimate missing data on non-completers, or include all subjects with pretests but not all who were randomized. These modified ITTs would be coded as “2” below. Completer analyses (AKA ‘per-protocol’, ‘efficacy’, or ‘exploratory’ analyses) involve only the subjects who stayed in the study, or only those who completed treatment.

1. Intent-to-treat analysis (all subjects who were assigned are used in posttest)
2. Modified intent-to-treat (not all assigned subjects are used in posttest, but authors have done some modifications to approximate a true ITT)
3. Completer analysis (only those subjects who completed treatment or who stayed in the study are used in posttest)

**Assigned and Observed N**

Assigned N, Observed N. These fields refer to the number of subjects who were originally assigned to the group(s) involved in this effect size (Assigned N) and to the number of subjects who were actually “observed” or “measured” (Observed N). If you cannot tell how many subjects were originally assigned to a group, look at the number of subjects (Observed N) at pretest; you can frequently use pretest sample sizes for assigned N. However, in cases where the authors have removed the subjects who do not have both pretest and posttest measures (such that the pretest N and the posttest N are the same), do not assume that the number of subjects at pretest is the
correct number for Assigned N and, instead, leave this field blank. In cases where there is no attrition, the Assigned N is the same as the Observed N. Only use the same numbers for Assigned N and Observed N when you are SURE that there is no attrition.

[ES36] Assigned N for the treatment group (or pretest, if this is a pretest-posttest effect size).
[ES37] Assigned N for the comparison or second treatment group (or posttest, if this is a pretest-posttest effect size; if this is a pretest-posttest effect size, this value should be the same as the assigned N for the pretest).
[ES38] Total Assigned N.

[ES1] Observed N for the treatment group (or pretest, if this is a pretest-posttest effect size).
[ES2] Observed N for the comparison or second treatment group (or posttest, if this is a pretest-posttest effect size).
[ES3] Total Observed N.

Other Effect Size Data Fields

[ES9] Mean for treatment group
[ES10] Mean for comparison group
[ES11] Difference in group means
[ES12] Standard deviation for treatment group
[ES13] Standard deviation for comparison group
[ES14] Pooled sd
[ES31] N successful for treatment group
[ES32] N successful for comparison group
[ES33] N failed for treatment group
[ES34] N failed for comparison group

[ES4] Dependent t-value
[ES5] Independent t-value
[ES6] \( \chi^2 \) (df=1)
[ES20] Effect size reported by authors
[ES60] Odds ratio reported by authors

Final Effect Size Determination

[ES21] Effect size value- standardized mean difference
[ES81] Effect size value- odds ratio

Remember that you cannot rely on simple numerical values to determine which group has done better. For treatment-control comparisons, a positive effect size
should indicate that the treatment group did “better” on the outcome measure than the comparison group, while a negative effect size indicates that the comparison group did “better” than the treatment group, and a zero effect size means that the two groups are exactly equal on the measure. For single-group pretest-posttest comparisons, a positive effect size indicates that the group did better at posttest than at pretest, while a negative effect size indicates that the group did better at pretest than at posttest, and a zero effect size means that the group’s performance was exactly equal at the two time points.

You must make sure that the sign of the effect size matches the way we think about direction, such that the effect size is positive when the treatment group (or posttest) is better and negative when the comparison group (or pretest) is better.

Effect sizes can range anywhere from around –3 to +3. However, you will most commonly see effect sizes in the –1 to +1 range.

Note: If the authors report an effect size, include that in your coding and use it for the final effect size value if no other information is reported. However, if the authors also include enough information to calculate the effect size, always calculate your own and report it in addition to that reported in the study.

[ES39] Any problems coding this effect size?
### References for Studies Included in the Meta-analyses

<table>
<thead>
<tr>
<th>StudyID</th>
<th>Citations</th>
</tr>
</thead>
</table>


<table>
<thead>
<tr>
<th>StudyID</th>
<th>Citations</th>
</tr>
</thead>
<tbody>
<tr>
<td>StudyID</td>
<td>Citations</td>
</tr>
<tr>
<td>---------</td>
<td>-----------</td>
</tr>
</tbody>
</table>


StudyID: Citations

Demonstration Research Corporation.


StudyID  Citations


StudyID  Citations
BEST on students' dropout rates. Los Angeles, CA: University of California, Los Angeles, National Center for Research on Evaluation, Standards and Student Testing (CRESST), Center for the Study of Evaluation (CSE), Graduate School of Education & Information Studies.


<table>
<thead>
<tr>
<th>StudyID</th>
<th>Citations</th>
</tr>
</thead>
</table>
StudyID | Citations
--- | ---
NY: Manpower Demonstration Research Corporation.

44 Community College of Rhode Island. (1994). *How to implement a tech prep program based on the Rhode Island model*. Warwick, RI: Community College of Rhode Island.


<table>
<thead>
<tr>
<th>StudyID</th>
<th>Citations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>18789</td>
</tr>
</tbody>
</table>


<table>
<thead>
<tr>
<th>StudyID</th>
<th>Citations</th>
</tr>
</thead>
</table>


Rumberger, R. W., & Larson, K. A. (1994). Keeping high-risk Chicano students...


Citations


<table>
<thead>
<tr>
<th>StudyID</th>
<th>Citations</th>
</tr>
</thead>
<tbody>
<tr>
<td>StudyID</td>
<td>Citations</td>
</tr>
<tr>
<td>---------</td>
<td>-----------</td>
</tr>
</tbody>
</table>


<table>
<thead>
<tr>
<th>StudyID</th>
<th>Citations</th>
</tr>
</thead>
</table>
StudyID  Citations


<table>
<thead>
<tr>
<th>StudyID</th>
<th>Citations</th>
</tr>
</thead>
</table>