Family Life Education
Guide to Conducting a Critical Review of the Research
Mary Kay Keller M.P.A., PhD
Chair of Applied Family Sciences Master’s Program
Wilmington University
Table of Contents

INTRODUCTION .................................................................................................................. 3

CRITERIA GUIDELINES ..................................................................................................... 3

I. Define The Question: ...................................................................................................... 3

II. Comprehensive Search Strategy: .................................................................................. 3

III. Methodology Of Selection, Quality Assessment, And Reporting Results: ............... 4

A. RESEARCH EVALUATION ............................................................................................ 4

   Criteria For Reviewing All Research Articles Methodology ...................................... 4
   Criteria For Reviewing Quantitative Research Articles. ............................................ 5
   Criteria For Reviewing Qualitative Research Articles .............................................. 9

B. ALTERNATE EVALUATIONS FOR BEST PRACTICES AND STRATEGIES 11

   Family Life Education Programs: ............................................................................. 11

REFERENCES .................................................................................................................. 15

Adapted from, Criteria determining rigor of evidenced based Substance Abuse Prevention Programs and Critical Literature Reviews, State of Florida Substance Abuse Program Prevention Office by Mary Kay Keller, MPA (2011)
INTRODUCTION

The purpose of a critical literature review is to conduct a rigorous search of current and past research, models, best practices and programs that target age appropriate people Adults (18-55+ years of age) and/or children (prenatal- 17 years of age). This rigorous search will result in a thorough knowledge base that summarizes relevant literature in the field of FLE (Family Life Education) and identifies strengths and weaknesses of each. This literature review will involve the critical evaluation of a wide range of important scientific and clinical publications relating to FLE.

CRITERIA GUIDELINES

The guidelines for this process will focus and evaluate the following: the definition of the question, the comprehensiveness of the search strategy, the methods of choosing and assessing the primary studies, and the methods of combining the results and reporting results of the review. The review will be written in language for the intelligent layperson, i.e. someone who may not have a detailed knowledge of the field.

I. Define The Question:

Aim of the literature to be reviewed:

- Developing a full and vigorous integration of two theoretical frameworks:
  EXAMPLES: Life Span Development and Ecological Theories.

- provide a context for describing, elaborating and evaluating FLE research, evidence based strategies, best practices, and model programs that target adults (18-25, 25-55, and 55+) and/or children (prenatal- 17 years of age).

II. Comprehensive Search Strategy:

- Peer reviewed journal articles: Fields of relevant study, if not available in NCFR (National Council of Family Relations Journals) and/or other disciplines focused on FLE such as (and not limited to) Family Studies, Psychology, Social Work, Health Sciences or Public Health are encouraged, and

---


2 Insert the identified Family Theory for aim of this critical review.

Adapted from, Criteria determining rigor of evidenced based Substance Abuse Prevention Programs and Critical Literature Reviews, State of Florida Substance Abuse Program Prevention Office by Mary Kay Keller, MPA (2011)
III. Methodology of Selection, Quality Assessment, and Reporting Results:

- Titles Concise and explicitly states the purposes and goal;
- Authors - The author/s may or may not be well known in the field, indicate the institution or project to which the author/s are affiliated or supported;
- Abstract: Concise and explicit statement of what, by or to whom and to what extent were the results;
- Introduction: Includes key concepts, relevant literature, aim, evaluation of aim achieved, explicit description of operational variables, research strategy definition, and rationale;
- Literature Review Report includes: focus on the research versus the researcher, description of the limitations of the research and how these limitations are addressed, and integration of the research to define the theories.

A. RESEARCH EVALUATION

Criteria for Reviewing All Research

1. **Who**

   The author describes who is treated by stating the subjects(s)’ average age and standard deviation of age, and sex or proportion of males and females, and clearly defines the behavior(s) to be treated.

---

3 Applies to Program Evaluation
4 Applies to Program Evaluation

Adapted from, *Criteria determining rigor of evidenced based Substance Abuse Prevention Programs and Critical Literature Reviews*, State of Florida Substance Abuse Program Prevention Office by Mary Kay Keller, MPA (2011)
2. **What**

   The authors describe the intervention so specifically that you could apply the program with nothing more to go on than their description or they refer you to a book, videotape, CCD_ROM, articles, or Web address that (provides explicit) description of the program.

3. **Where**

   Authors state where the program occurred so specifically that you could contact people who conducted the intervention by phone, letter, or E-mail address.

4. **When**

   Authors tell the when of the intervention by stating how long subjects participated in it in days, weeks, or months or tell how many treatment sessions subjects attended.

5. **Why**

   Authors discuss a specific theory that describes why the intervention should work, or they cite literature research that demonstrates the intervention’s effectiveness.

   *Quantitative Research Articles proceed to #6 below.*

   *Qualitative Research articles skip the questions below; proceed to page 9 to #6.*

   Criteria for Reviewing Quantitative Research Articles.

6. **Subjects randomly assigned to intervention or control**

   The author states specifically those subjects were randomly assigned to the intervention groups or were referred to the assignment of subjects based on a table of random numbers, computer algorithm, or accepted randomization procedure. This means that the procedure resulted in each subject having an equal chance of being assigned to the intervention or control groups. Random assignment ensures better than any other procedure that control or program groups are initially similar before treatment begins, so post treatment differences can be attributed to effects of the program. Random assignment concerns the internal validity of a study. If the author says subjects were randomly assigned but assigns subjects by assigning every other one from a list or by allowing subjects or others to choose the treatment groups somehow other than by random procedure, then subjects are not randomly assigned.
7. **Analysis shows equal program and control groups before treatment**

Even though subjects have been randomly assigned, unequal intervention and control groups can occur by chance. To guard against initial dissimilarity, the authors need to make comparisons across the intervention group and control groups on key client characteristics to see that they are similar prior to treatment (e.g., risk factors related to the behavior to be prevented).

8. **Subjects blind to being in intervention or control group**

Subjects who know they are in a control group can experience effects of being there including demoralization or competition with experimental (Cook 7 Campbell, 1979, p. 55). Subjects who know they are in an intervention group can experience positive effects because they expect those (Brown, 1998). Give points for subjects blinded if two or more groups get some kind of intervention, if controls get some form of a program that is not expected to have an effect but gives assurance that something is being done, or if subjects serve in a delayed intervention control group where they serve as controls but get the intervention later, or if subjects truly do not know whether they are in a intervention group or control group.

9. **Subjects randomly selected for inclusion in study**

*Selection* of subjects is different from *random assignment*. Random selection means subjects are taken from some potential pool of subjects for inclusion in the study by using a table of random numbers or other statistically random procedures. For example, if subjects are chosen randomly from among all high-risk teenagers in a school, the results of the study can be generalized more confidently to all such students in that school. Random selection concerns generalizing results of the study to others, or external validity.

10. **Control (non-treated) group used**

Members of a *non-treated control group* do not receive a different kind of intervention; they receive *no* treatment. An example of a non-treated control group would be a group of teenage girls at high risk for pregnancy who do not get into a pregnancy-prevention program, while others are given an explicit pregnancy-prevention program. Subjects in the non-treated control group might receive the program at a later date but do not receive it while program group subjects are receiving their program.

Adapted from, *Criteria determining rigor of evidenced based Substance Abuse Prevention Programs and Critical Literature Reviews*, State of Florida Substance Abuse Program Prevention Office by Mary Kay Keller, MPA (2011)
11. **Number of subjects in smallest intervention group exceeds 20**

    Those in the intervention group or groups are those who receive some kind of special care intended to help them. It is this treatment (intervention program that is being evaluated by those doing the study. The results of the study will state how effective the intervention programs have been when compared with each other or with a non-treated control group. In order to meet Criterion 11, *the number of subjects in the smallest (intervention) group must be at least 21.* (Not everyone will agree with this number. Apply a statistical power analysis if possible.) Here, *number of subjects* means total number of individuals, not number of couples or number of groups.

12. **Outcome measure has face validity**

    Face validity is present if the outcome measure used to determine the effectiveness of treatment makes sense to you. Regarding intervention programs, such outcomes should truly reflect what the intervention program intends to accomplish. A good criterion for the sense of an outcome measure is whether the measure evaluates something that should logically be affected by the program. For example, a school violence-prevention program should reduce the number of reported incidents of physical fights among students. This would be more valid as an outcome than the number of violence-prevention meetings attended.

13. **Treatment outcome measure was checked for reliability**

    For this criterion to be met, it is not enough to merely say that the outcome of the intervention was measured in some way. The outcome measure itself must be evaluated to check its reliability. *Reliability-* the consistency of measurement- is frequently measured in an outcome evaluation study by comparing the findings of investigators who independently rate the performance of individuals in treatment or control groups. Another less frequently used way to measure reliability of outcome measures is to have the same individual rate the performance of subjects, then re-rate their performance. In single-subject studies, two raters may rate the subject’s behavior independently for cross-rater comparison.

    The reliability criterion is satisfied only if the author of the study affirms that evaluations were made of the outcome measure’s reliability (for example, inter-rater agreement), and *the author lists a numerical value of some kind for this measure of reliability.* Where multiple

---

Adapted from, *Criteria determining rigor of evidenced based Substance Abuse Prevention Programs and Critical Literature Reviews*, State of Florida Substance Abuse Program Prevention Office by Mary Kay Keller, MPA (2011)
outcome criteria are used, reliability checks of any one of the major outcome criteria satisfy criterion 13

14. *Reliability measure has value greater than* .70 *or percent of rater agreement greater than* 70%

   The reliability coefficient (alpha) in Criterion 13 is .70 or greater (70% or better).

15. *Those rating outcome rated it blind.*

   This criterion concerns the way bias can center into measurement if the person measuring the outcome knows whether the subject being measured is from a treatment or control group, or, worse, the person measuring the outcome is in a position to determine the outcome measure. As an example for the latter, a large study had as an outcome whether chronically mentally ill persons treated in the community were readmitted to a hospital. Those administering the special treatment were in a position to decide whether the clients would be readmitted, which was an outcome criterion in the study (Test 7 Stein, 1977, pp.14-15; Gomory, 1999, pp. 154-155). *Give the points for this criterion only if the person conducting the outcome measuring did not know which subjects were in the intervention or control groups.*

16. *Outcome of the intervention was measured after the program was completed*

   At least one outcome measure was obtained after the intervention was completed. Ideally, an intervention’s effects should not decay to produce no effect after the program ends. Its effects should to continue for some interval after the program ends. For example, if high risk teenagers receive an intervention that ends at the end of the school year in 9th grade, then its effects should be observed in the grades to follow. If outcome is measured both during and after the intervention program, it also meets this criterion.

17. *Test of statistical significance was made and p<.05*

   Tests of statistical significance are generally referred to by phrases such as differences between intervention and control groups were significant at the .05 level” or “results show statistical significance for…. “ Give credit for meeting this criterion only if the author refers to a test of statistical significance for a major outcome variable stating the name of the statistical test (e.g., analysis of variance, chi square, t test) and gives a p value, for example p<.05, and the p value is equal to or smaller than .05. (With small samples in exploratory

Adapted from, *Criteria determining rigor of evidenced based Substance Abuse Prevention Programs and Critical Literature Reviews*, State of Florida Substance Abuse Program Prevention Office by Mary Kay Keller, MPA (2011)
studies, .10 may be better to reduce the chance that one will miss a real difference, but generally .05 is the convention).

18. **Follow-up was greater than 75%**.

The proportion of subjects successfully followed up refers to the number contacted to measure outcome compared with the number who began the study. Ideally, the two should be the same. To compare the proportion followed up for each group studied (e.g., intervention group(s), control group), determine the number of subjects who initially entered the groups and determine the number successfully followed up. (If there is more than one follow-up period, use the longest one.) Then, for each group, divide the number successfully followed up by the number who began in each group and multiply each quotient by 100. For example, if 20 entered an intervention group but 15 were followed up in that group, the result would be: \((15/20) \times 100 = 75\%\). Compute the proportion followed up for all groups involved in the experiment. If the smallest of these percentages exceeds 75%, the study meets this criterion.

19. **Base rate comparison** has particular importance as a standard for judging intervention programs.

Ideally, careful records within the agency will show the prior rate of the behavior before the intervention program began. This base rate experience can provide a benchmark to judge the effects of the program. For example, the rate of pregnancy for the high-school juniors and seniors 2 years prior to the program can be compared with the rate of juniors and seniors during the 2 years after the program. **Give points for this criterion only if records have been kept for a specific interval (e.g., two years) regarding the rate of the behavior among high-risk persons prior to the intervention program and also during the same interval of time after the intervention program, and the behavior changes substantially- you decide what substantially means.**

**Criteria for Reviewing Qualitative Research Articles**

Qualitative research is used to seek understanding about participants' beliefs and experiences anchored in the world we most intimately know (Freeman, 2004, p. 79). Qualitative methodology provides the opportunity to capture and highlight the nuances of individual responses of participants in a study.

Adapted from, *Criteria determining rigor of evidenced based Substance Abuse Prevention Programs and Critical Literature Reviews*, State of Florida Substance Abuse Program Prevention Office by Mary Kay Keller, MPA (2011)
6. Credibility (*Internal Validity)*:

Did the research include descriptions of triangulation of the data, member checking and time sampling? All are examples of rigorous qualitative methodology.

*Triangulation*:

Did the research methodology include multiple data sources? Triangulation may include data sources such as demographics, structured interviews, diaries, self reports, pictorial data (video or pictures), researcher self reflection diaries, etc. This builds trustworthiness and rigor of the research (Patton, 2002).

*Member Checking*:

Were participants asked to review their responses to interviews and other data collection they participated in for accuracy of the data? If there are any discrepancies member checking allows the data to be corrected to best represent the responses of the participants. Member checking provides credibility (accuracy) the process of analysis by controlling for researcher bias and correcting any misunderstanding of the participants recorded responses (Saldana, 2001). This supports accuracy and validation of the findings.

*Time Sampling*:

In qualitative research it is often necessary to collect data over time, interval time sampling. Did the article describe in detail when the data was collected? Was there a system by which the time was determined to be appropriate? Was that system described? If there was just a pre and a post collection, how many days, weeks, months or years were in between the pre and the post collection? Was a rationale included in the article?

7. Transferability (*External Validity)*:

Provide thick description

Did the research article provide precise descriptions of the research design, methodology and data reports? How were the participants selected? Did the article clearly describe the process in which the participants were selected? It is common for qualitative researchers to use small, purposive samples to capture unique details and noteworthy meanings (Berg, 2007).

8. Dependability (*Reliability)*:

*Audit Trail*:

Did the research indicate there were coding and data memo’s describing the data analysis and was the data analysis described in the article? It is imperative that data memo’s and the

Adapted from, *Criteria determining rigor of evidenced based Substance Abuse Prevention Programs and Critical Literature Reviews*, State of Florida Substance Abuse Program Prevention Office by Mary Kay Keller, MPA (2011)
researchers own reflexivity journal are kept in detail to describe the procedures so that the research can be replicated by others.

*Code-Recode Strategy:*

Was the coding and recode strategy described in detail in the memos. Were the procedures for coding and recoding described included in the article or referenced in the article?

*Triangulation:*

Did the research methodology include multiple data sources? Triangulation may include data sources such as demographics, structured interviews, diaries, self reports, pictorial data (video or pictures), researcher self reflection diaries, etc. This builds trustworthiness and rigor of the research (Patton, 2002).

9. **Confirmability (Objectivity):** Triangulation, Reflexivity, Negative Case Analysis

*Triangulation:*

Did the research methodology include multiple data sources? Triangulation may include data sources such as demographics, structured interviews, diaries, self reports, pictorial data (video or pictures), researcher self reflection diaries, etc. This builds trustworthiness and rigor of the research (Patton, 2002).

*Reflexivity:*

Were there research Memos documenting all decisions in detail? Was the research position, attitude and perspective disclosed in the research or indicated a presence in the article?

*Negative case analysis:*

Were there examples of data that did not support the outcomes? Full disclosure is imperative in qualitative research to address experimenter bias. Data that does not support the outcomes has to be examined as to whether or not it negates the outcomes, is non-essential to the outcomes or is of standalone importance (Anfara et al., 2002; Patton, 2002).

B. ALTERNATE EVALUATIONS FOR BEST PRACTICES AND STRATEGIES

Family Life Education Programs:

1. **Family and Community Involvement**

Programs that include a commitment from communities, families and school districts have shown much higher success rates in their prevention.

---


Adapted from, *Criteria determining rigor of evidenced based Substance Abuse Prevention Programs and Critical Literature Reviews*, State of Florida Substance Abuse Program Prevention Office by Mary Kay Keller, MPA (2011)
Supporting Citation:

2. **Comprehensive, Multi-Component**

Programs that use a combination of (1) normative education, (2) information about the outcomes and (3) social skills training, including social influences training (especially peer pressure resistance skills) are more successful than using a single approach.

Supporting Citations:


3. **Information Dissemination**

The interactive process provides opportunities for the exchange of ideas between peers. Researchers found that non-interactive programs showed only a 4% reduction in (substance use) prevalence rate, while Interactive programs showed a 21% reduction in prevalence rate. Small interactive programs were found to be most successful.

Supporting Citation:

4. **Interactive Teaching Techniques** (1 pt.)

The interactive process provides opportunities for the exchange of ideas between peers. Researchers found that non-interactive programs showed only a 4% reduction in behaviors prevalence rate, while Interactive programs showed a 21% reduction in behavior prevalence rate. Small interactive programs were found to be most successful.

Supporting Citation:

5. **Long-term, Multi-Year Programming** (5 pts.)

Studies have found that in order for an education program to be effective, it needs to be delivered over a long period of time to continually reinforce skills.

Supporting Citations:

Adapted from, *Criteria determining rigor of evidenced based Substance Abuse Prevention Programs and Critical Literature Reviews*, State of Florida Substance Abuse Program Prevention Office by Mary Kay Keller, MPA (2011)
6. **Mentoring** (0 pts.)

Mentoring is one-on-one interaction with an older, more experienced person to provide advice or assistance. Evaluations of community-based mentoring programs have found the programs decrease substance use. School-based mentoring programs show promise in that they reinforce strengths that may lead to the delayed onset of negative behaviors or participation in negative behavior.

Supporting Citation:

7. **Normative Education** (5 pts.)

Clarifying and Communicating norms about behavior. An essential part of the Comprehensive, Multi-Component Approach.

Supporting Citations:


8. **Parent Involvement** (0 pts.)

The participation of parents in a prevention program has been found to help increase communication, alter students’ attitudes toward positive health practices

Supporting Citation:

9. **Refusal/Resistance Skills Training** (3 pts.)

Activities that teach refusal or resistance skills are incorporated into the program along with opportunities for practice. These programs help prepare students to identify pressures to use drugs and give students the skills they need to resist peer pressure to use drugs.

Adapted from, *Criteria determining rigor of evidenced based Substance Abuse Prevention Programs and Critical Literature Reviews*, State of Florida Substance Abuse Program Prevention Office by Mary Kay Keller, MPA (2011)
Supporting Citations:


10. **Social Influences** (5 pts.)
An emphasis on social influences such as advertising and media as well as the influence of friends (peer resistance skills training) and family members as role models are an important part of the **Comprehensive, Multi-Component Approach**. Research has shown that a focus on social influences is a critical aspect of effective drug prevention education.

Supporting Citation:

11. **Social Skills Training** (3 pts.)
Social Skills Training means focusing on a range of social competency skills (e.g. developing self-control, stress management, responsible decision-making, social problem solving, and communication skills). It is an integral part of the **Comprehensive, Multi-Component Approach**.

Supporting Citations:

Adapted from, *Criteria determining rigor of evidenced based Substance Abuse Prevention Programs and Critical Literature Reviews*, State of Florida Substance Abuse Program Prevention Office by Mary Kay Keller, MPA (2011)
REFERENCES


Adapted from, *Criteria determining rigor of evidenced based Substance Abuse Prevention Programs and Critical Literature Reviews*, State of Florida Substance Abuse Program Prevention Office by Mary Kay Keller, MPA (2011)


Adapted from, *Criteria determining rigor of evidenced based Substance Abuse Prevention Programs and Critical Literature Reviews*, State of Florida Substance Abuse Program Prevention Office by Mary Kay Keller, MPA (2011)