

How Were Reviews of the Research on the *Elements of Effective Practice*[™] Conducted?

The reviews of the evidence relating each area (Standard and associated Benchmarks and Enhancements) of the *Elements of Effective Practice – Third Edition*[™] were conducted by the Research Board Chair of the National Mentoring Resource Center (Dr. David DuBois) and two postdoctoral research associates working under his supervision (Drs. Kevin Pinkston and Gizem Erdem).

The first step in conducting each review was to systematically and comprehensively search for relevant studies. Key search strategies include:

- Keyword-based searches in on-line databases of published research for relevant disciplines such as psychology, education, social work, criminal justice, and public health;
- Review of the numerous publications on mentoring produced by Public/Private Ventures;
- Review of presentations made at the National Mentoring Summit;
- Review of literature cited in previous overviews or syntheses of research relating to the Standard as well as in the relevant section of the *Elements*;
- Contacting the youth mentoring listserv, NMRC Research Board members; and
- Review of literature cited in studies identified through the above literature search methods.

The next step was to assess each identified study against criteria for inclusion in the review. To be included, a study needed to report findings that were directly pertinent to the Standard being reviewed. In particular, the study needed to examine the implications of one or more practices relating to the Standard for program effectiveness (e.g., youth outcomes, mentoring relationship quality), efficiency (e.g., mentor retention), sustainability (e.g., funding), or safety (e.g., incidents of harm to youth). If study was quantitative in its design, it needed to either compare measured outcomes in one of these areas across programs or participants who were and were not exposed to the practice or, alternatively, change in outcomes before and after exposure to the practice for a single group of programs or participants. These types of studies are encompassed by Evidence Levels 1a and 1b in a framework characterizing evidence for mentoring program practices that was developed by Dr. DuBois (see http://www.mentoring.org/downloads/mentoring_383.pdf). Qualitative studies (those that collect and analyze non-quantitative data such as those obtained in focus groups or interviews) could entail comparisons across different programs or participants as well as case studies of a single program or participant (Evidence Levels 3a and 3b in the above framework).

After preparing a written summary of the findings of each identified study, the final step in each review was to rate the available evidence along the following 4 dimensions:

- **Extent of Evidence Base:** Number of studies
- **Favorability of Findings:** Degree to which findings are supportive of practices that align with the Standard
- **Methodological Rigor:** Strength of research design for reliably and validly detecting effects of the relevant program practices
- **Scope of Findings:** Degree to which findings address implications of multiple practices that align with the Standard as well as use of relevant practices within different program models (e.g., site- and community-based) and with mentors and youth of varying backgrounds (e.g., older peer and adult mentors, youth faced with differing types of life challenges).

The evidence for each dimension was rated on a 4-point scale (see “Evidence Rating Scales for Reviews of Elements of Effective Practice”). Initial ratings were made independently by Dr. DuBois and each of the postdoctoral research associates. Following a conference, each reviewer made his or her final ratings and these were then averaged across reviewers.

Evidence Rating Scales for Reviews of *Elements of Effective Practice*[™]

Extent of Evidence Base: Number of studies.

3 - High: More than 10 eligible studies, where multiple reports based on data from the same investigation are considered to constitute one study.

2 - Moderate: 6-10 eligible studies

1 - Low: 1-5 eligible studies

0 - None: No eligible studies

Favorability of Findings: Degree to which findings are supportive of practices that align with the Standard.

3 - High: Findings consistently indicate benefits of practices that align with the Standard, with exceptions limited to a small number of non-significant results and none that suggest adverse effects.

2 - Moderate: Findings as a whole support benefits of practices that align with the Standard, but with a substantial number of non-significant findings and potentially some limited evidence of possible adverse effects.

1 - Limited: Findings as a whole do not indicate benefits of practices aligned with the Standard, although there are some specific results that are supportive; there may also be limited evidence of adverse effects of the practices of interest.

0 - None: Findings include no (or very limited) evidence of benefits of practices aligned with the Standard or there is substantial evidence of adverse effects of such practices.

Methodological Rigor: Strength of research design for reliably and validly detecting effects of the relevant program practices.

3 - High: Evidence base includes either:

Multiple, rigorously-conducted random-assignment studies

OR

At least one rigorous random-assignment study in combination with at least 2 studies that are quasi-experimental in design (i.e., do not use random assignment, but do include a comparison group of programs or participants not receiving the practice) and methodologically strong overall (i.e., rigorously designed and executed).

2 - Moderate: Evidence base includes either:

Multiple, methodologically-strong quasi-experimental studies

OR

One methodologically-strong experimental or quasi-experimental study in combination with either multiple quasi-experimental or experimental studies that do not meet standards for highest rigor or multiple, methodologically-strong qualitative studies (i.e., studies include multiple programs or participants and are rigorously designed and executed)

1 - Low: Evidence base includes either:

Experimental or quasi-experimental studies, but these do not meet highest standards for rigor

OR

Exclusively qualitative studies

0 - Very Low: Evidence base does not include studies satisfying criteria for any of the above ratings.

Scope of Findings: Degree to which findings address implications of multiple practices that align with the Standard as well as use of relevant practices within different program models (e.g., site- and community-based) and with mentors and youth of varying backgrounds (e.g., older peer and adult mentors, youth faced with differing types of life challenges). Ratings are made according to the how many of these 4 dimensions are addressed; for a dimension to be considered as addressed, findings must address effects of three (3) or more distinct practices (e.g., different Benchmarks or Enhancements associated with the Standard), program models (e.g., site-based, community-based, group), or distinct groups of mentors (e.g., adult mentors, older peer mentors, paid mentors) or youth (e.g., children, adolescents, youth in the child welfare system), respectively.

3 - Extensive: Findings address at least 3 of the 4 dimensions of scope according to the above criteria.

2 - Moderate: Findings address 2 of the 4 dimensions.

1 - Limited: Findings address only 1 of the 4 dimensions.

0 - Very Limited: Findings do not address any of the dimensions of scope.