ARTICLE

# RESEARCH METHODOLOGY AND YOUTH MENTORING

David L. DuBois
*University of Illinois at Chicago*

Fred Doolittle
*Manpower Demonstration Research Corporation (MDRC)*

Brian T. Yates
*American University*

Naida Silverthorn
*University of Illinois at Chicago*

Jacob Kraemer Tebes
*Yale University School of Medicine*

*Mentoring programs for youth have grown tremendously in popularity in recent years and in many important respects reflect core principles of community psychology. Mentoring of youth is a complex phenomenon, however, with a range of significant processes occurring at the levels of individual youth and their mentors, youth–mentor relationships and other interpersonal systems, programs, and the larger policy context. The research methods used to study youth mentoring need to be well suited to capturing this complexity. In this article, we argue, furthermore, that investigations of youth mentoring relationships and programs should be tailored to address concerns associated with each major phase of the intervention research cycle (i.e., preintervention, intervention, and preventive service systems research). Existing research pertinent to these differing phases frequently has not employed state-of-the-art methodology in the areas of sampling, design, assessment, and analysis. We also find that there are important gaps in the types of research conducted, and that in many instances, needed linkages across phases of the research cycle are lacking. Recommendations for strengthening future research on youth mentoring are discussed.* © *2006 Wiley Periodicals, Inc.*

The writing of this article was supported in part by a grant to the first author from the National Institute of Mental Health (1 R21 MH069564).
Correspondence to: David L. DuBois, Institute for Health Research and Policy, University of Illinois at Chicago, 1747 W. Roosevelt Rd., Chicago, IL 60608. E-mail: dldubois@uic.edu

**JOURNAL OF COMMUNITY PSYCHOLOGY**, Vol. 34, No. 6, 657–676 (2006)
Published online in Wiley InterScience (www.interscience.wiley.com).
 DOI: 10.1002/jcop.20122

Historically, mentoring programs for youth emerged from grassroots efforts of social activists (Baker & Maguire, 2005). Today, it continues to be the case that the vast majority of mentoring programs for young people originate in community settings and are operated by practitioners (DuBois & Karcher, 2005). These programs tend to reflect many of the values that are embraced most closely by the field of community psychology (Dalton, Elias, & Wandersman, 2001), including citizen participation (via use of community volunteers as mentors), respect for human diversity (via cultural tailoring of programs to minority and other diverse youth populations), and an emphasis on community strengths (via utilization of existing youth-serving agencies and organizations as sites for program development and implementation). Few existing programs, however, have benefited from development and evaluation within empirically driven frameworks (DuBois & Silverthorn, 2005a. Consequently, although mentoring initiatives for youth have soared in popularity in recent years and now number in the thousands (Rhodes, 2002), the advancement of a strong empirical grounding for these initiatives has lagged significantly behind (DuBois & Karcher, 2005).

Mentoring of youth is a complex phenomenon with a range of important processes occurring at the levels of individual youth and their mentors, youth–mentor relationships and other interpersonal systems, programs, and the larger policy context. The methods used to study mentoring of youth need to be well suited to capturing this complexity (DuBois & Silverthorn, 2005a). In this article, we argue, furthermore, that research efforts should be distributed and integrated across the different phases of activity that have been viewed as essential for ultimately achieving large-scale, community- and population-level impacts on targeted outcomes (Flay, 1986; Institute of Medicine [IOM], 1994; National Advisory Mental Health Council Workgroup on Mental Disorders Prevention Research [NAMHC], 2001). These phases, drawn from the National Institute of Mental Health (NIMH, 1998), include preintervention research, preventive intervention research, and preventive service systems research (see Figure 1). Preintervention research may help to identify fundamental mechanisms influencing the development and maintenance of mentoring relationships and their consequences for health and well-being. It also provides an opportunity to conduct research focused on development and preliminary analysis of new intervention strategies. Building on this foundation of knowledge, intervention research then can investigate the efficacy and effectiveness of promising mentoring programs and initiatives. Finally, for those approaches found to be beneficial in well-controlled studies, research can assume an integral role in identifying effective approaches to their dissemination within the broader preventive service systems and policy context. As illustrated in Figure 1, the research cycle is assumed also to include reciprocal linkages.

This article reviews methodology in the study of mentoring with respect to ways it may be used to advance state-of-the-art research in each of the major phases of an empirically driven framework for intervention development, evaluation, and dissemination (i.e., preintervention research, intervention research, and preventive service systems research). In accordance with an action research perspective (Dalton et al., 2001), we emphasize the importance of collaboration with community partners at each stage of the research process.

## PREINTERVENTION RESEARCH

Preintervention research is widely regarded as an essential step in the development of any preventive intervention (Flay, 1986; IOM, 1994; NAMHC, 2001; NIMH, 1998). In the context of mentoring research, preintervention investigations can be conceptualized as

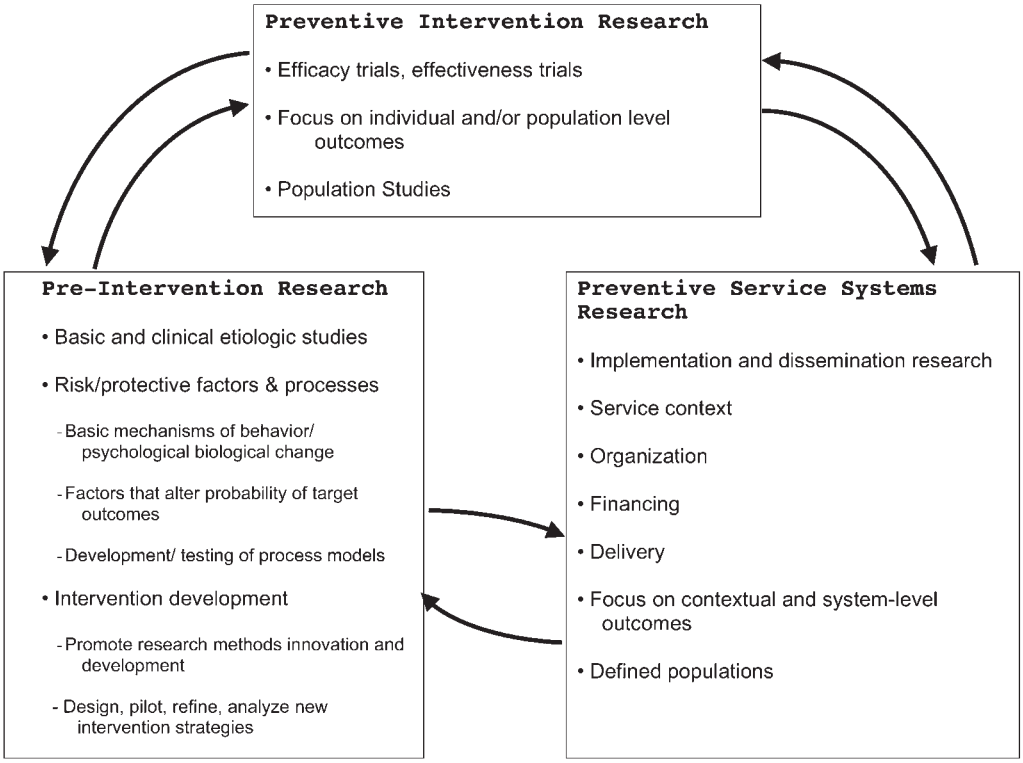


**Figure 1.** Comprehensive prevention research cycle.

serving two primary purposes. First, they offer an opportunity to conduct basic research that helps to delineate those conditions and processes that serve both to maximize the potential for youth to accrue positive developmental gains from their involvement in mentoring relationships and programs and to minimize the risk for youth to be harmed by these experiences (DuBois & Karcher, 2005). These insights then can be incorporated into the design of intervention strategies, improving the odds that they will yield positive and substantial benefits for youth when subjected to rigorous testing in the next phase of the research cycle. Second, preintervention studies provide an opportunity to design, pilot, refine, and conduct preliminary analyses of new intervention strategies within the mentoring field. These studies can be thought of as a bridge of sorts in the process of translating findings from basic research into actual strategies for intervention. Again, the assumption is that this work will pay dividends when interventions are subsequently evaluated for impact and cost-effectiveness in efficacy and effectiveness trials.

### *Basic Research*

Theoretically, a wide range of conditions and processes should be important in mediating and moderating the impact of mentoring relationships on youth outcomes (Rhodes, 2002, 2005; see also Keller, 2005; Sipe, 2005; Spencer & Rhodes, 2005). These include, but are not limited to, (1) attributes that the mentor and youth each take to the relationship, such as the mentor's skills and confidence and the youth's relationship history and current level of

functioning; (2) characteristics of the relationship, such as the extent to which mentor and youth form an emotional bond characterized by feelings of trust, empathy, and positive regard; the frequency and pattern of their contacts; the types of activities and discussions in which they engage; the ways in which needs for attention to both relationship development and instrumental, goal-focused concerns are integrated and balanced; the degree to which the mentor serves as a role model and advocate for the youth; and the relationship's duration; and (3) contextual factors, such as the preexisting network linkages to other important persons and relationships in the lives of both the youth and mentor and the characteristics of the program or other settings in which mentoring takes place. Likewise, although largely neglected to date, the types and value of resources used to provide mentoring are also multidetermined and must be elucidated in order to conduct cost-effectiveness and cost-benefit analyses and accurately gauge the potential cost-saving benefits of mentoring to social and health service (Yates, 2005). Creating order and understanding in the complex array of potential influences on both the benefits and costs of mentoring is no small task. The inroads made thus far, although noteworthy in several respects, are limited and incomplete. Methodological considerations relating to sampling, study design, assessment, and data analysis need to be addressed to begin to fill existing gaps.

*Sampling.* To date, most studies of mentoring relationships for youth have been based on relatively small samples of convenience. The size of the samples poses at least two significant problems. First, investigations have tended to lack adequate statistical power for detecting what often may be relatively subtle dynamics of mentoring relationships and their effects on youth outcomes (Keller, 2005; Rhodes, 2002, 2005). Second, it has not been possible to generalize findings to larger populations of interest with confidence.

There have been a few noteworthy exceptions to this trend. Two surveys of nationally representative samples of adults have asked them about their mentoring relationships with youth (AOL Time Warner Foundation, 2002; McLearn, Colasanto, & Schoen, 1998). These studies, however, failed to assess mentoring relationships from the perspective of youth. In the most recent wave of the Add Health Study (Bearman, Jones, & Udry, 2003), a longitudinal study of a nationally representative sample of adolescents, respondents (aged 18–26) were asked questions about mentoring relationships they had experienced since the age of 14 (for studies reporting on these data, see DuBois & Silverthorn, 2005b, 2005c). These data are limited, however, by their retrospective nature and the reliance on a single occasion of assessment (see the discussion of study design issues later). It also is noteworthy that none of the studies based on large, nationally representative samples has used a well-validated instrument to assess the features of mentoring relationships (see the discussion of measurement issues).

*Design.* From a design standpoint, studies of mentoring relationships for youth have been predominantly cross-sectional. Most investigations thus have not been well suited to addressing patterns of development and change in relationships over time that may have important implications for youth outcomes (Keller, 2005). Those longitudinal studies conducted, furthermore, typically have included only a limited number of assessments, which occurred over a relatively brief interval. Such designs lack the repeated observations necessary for refined assessment of the growth and evolution of relationships over their entire life course. Because of the time frames involved, the longer-term consequences of mentoring relationships also remain largely unexplored. Testimonials on behalf of mentoring frequently allude to its capacity to have enduring and transforma-

tive effects on youth that reach well into adulthood. There is a need, however, for empirical data that would allow for a rigorous evaluation of this assumption and the conditions under which it is most likely to hold true (Rhodes & DuBois, 2004).

*Assessment.* Paper-and-pencil questionnaire measures completed by youth and, in some instances, also by mentors have been the predominant approach to assessing relationships. Several instruments have been developed within the context of different research projects (for a review, see Nakkula & Harris, 2005). These measures are only beginning to be the focus of programmatic validation research. As a result, neither the psychometric properties of proposed instruments nor their appropriateness for use with different populations and types of mentoring is well established. Because of sampling issues discussed earlier, population-based normative data that would facilitate interpretation of relationship assessments and allow for meaningful comparisons across studies are not available.

A multimethod approach is essential both for validating measurements and for taking optimal advantage of state-of-the-art, multivariate data analytic procedures when investigating how mentoring relationships develop or influence youth (see the discussion of data analysis issues). It is thus noteworthy that few studies have incorporated the perspectives of both mentors and youth on relationships (for exceptions, see Parra, DuBois, Neville, Pugh-Lilly, & Povinelli, 2002; Karcher, Nakkula, & Harris, 2005). Nor have alternatives to questionnaire-based methods of assessing mentoring relationships, such as direct observation, received much attention (for an exception, see Newman, Morris, & Streetman, 1999).

It is of further note that most research has focused on the characteristics of a single mentoring relationship that has been either identified by the youth (in the case of studies of naturally occurring mentoring ties) or established by a program (in the case of studies of mentoring relationships within formal programs). Many youth, however, may experience a network of significant mentoring ties with multiple persons. Very little attention has been given to study of either the size or other potentially significant characteristics of these networks (e.g., linkages between mentors). Linkages of the mentoring relationship with other persons in the social networks of the youth and mentor (e.g., the youth's parent and teachers, members of the mentor's family) also are theoretically important but rarely have been investigated. Finally, assessment of the costs and potential monetary benefits entailed in mentoring relationships clearly would be valuable but has been exceptionally rare.

*Data analysis.* Data analyses in investigations of mentoring relationships often have been limited to bivariate procedures, such as zero-order correlations or *t* tests that compare measures across groups. More sophisticated, multivariate analyses, however, may be essential for addressing a variety of important concerns. Procedures, such as factor analysis and cluster analysis, are well suited to the task of identifying salient dimensions or typologies of mentoring relationships (Sipe, in press), for example, but have been employed only to a limited extent (e.g., Darling, Hamilton, Toyokawa, & Matsuda, 2002; Langhout, Rhodes, & Osborne, 2004; Liang, Tracy, Taylor, & Williams, 2002). Similarly, structural equation modeling provides a valuable, but underutilized technique for testing theories of mentoring and its effects on youth outcomes (e.g., DuBois, Neville, Parra, & Pugh-Lilly, 2002; Parra et al., 2002; Rhodes, Grossman, & Resch, 2000).

Because of the dyadic nature of relationships, and hence the contributions of both mentor and youth to their development, it also is important that data analytic techniques address processes that occur at this level. Few, if any, studies, however, have made use of relevant procedures that have been developed for the analysis of dyadic data (DuBois &

Silverthorn, 2005a). Likewise, there is growing interest in mentoring that occurs in a group context (Herrera, Vang, & Gale, 2002), as well as within particular types of settings, such as schools (Portwood & Ayers, 2005), the workplace (Hamilton & Hamilton, 2005), and after-school programs (Hirsch, 2005; Hirsch & Wong, 2005). Along with the considerations related to social networks, these developments underscore a need for analytic approaches that are sensitive to detecting how mentoring relationships may both be shaped by and shape features of the settings and environments in which they occur. Multilevel modeling procedures (e.g., hierarchical linear modeling) could be especially useful in this regard but apparently have not yet been employed for this purpose in the mentoring literature. Finally, analyses of costs and benefits of mentoring from the perspectives of participants in relationships and relevant others (e.g., parents) are conceptually feasible within naturalistic (preintervention) studies and could prove highly informative but, to our knowledge, have not been attempted.

### *Intervention Development*

The process of translating findings from basic research into intervention strategies and preparing for evaluation of these strategies within controlled trials requires attention to several different concerns (Bartholomew, Parcel, Kok, & Gottlieb, 2001; Flay, 1986; Green & Kreuter, 1999; IOM, 1994; NAMHC, 2001; NIMH, 1998). These include (1) utilizing the preferences and insights of stakeholder groups to inform the design of proposed intervention strategies, (2) piloting and refining intervention strategies before implementation of the intervention as part of a full-scale efficacy trial, and (3) developing a methodologically sound set of procedures for evaluating the intervention in terms of its effectiveness, benefits, costs, cost-effectiveness, and cost-benefit ratio. By addressing these concerns, preintervention research can enhance both the quality of mentoring interventions and the methodological rigor of the procedures that are used to evaluate them.

Stakeholders in mentoring interventions include the targeted population of youth and their caregivers, prospective volunteer mentors, staff and administrators of the mentoring agencies that will implement the intervention, and representatives of broader concerned entities, such as community coalitions, governmental agencies, and policy-making or advocacy organizations (Coyne, Duffy, & Wandersman, 2005). Input from these groups may be obtained through a variety of methodologies, including key informant interviews, focus groups, and surveys. As an illustration, the first author currently is obtaining input from multiple stakeholder groups as part of the process of further developing a mentoring program for young adolescent girls in collaboration with Big Brothers Big Sisters of Metropolitan Chicago. Key informant interviews are being conducted with girls and mentors who participated in an earlier version of the program, girls' parents, and staff and administrators of several Big Brothers Big Sisters (BBBS) agencies. To provide a mechanism for ongoing stakeholder input throughout the intervention development process, individuals from each stakeholder group are being recruited to serve on an advisory council.

Piloting the proposed components of interventions provides a further opportunity to assess their acceptability to stakeholders and, importantly, to refine components on the basis of lessons learned in the implementation process (Bartholomew et al., 2001; Flay, 1986). For mentoring programs, this could entail piloting procedures for key components of programs, such as mentor recruitment and screening, matching of youth with mentors, training and orientation, supervision, any special activities or services to be provided for mentors and youth, and procedures for monitoring the quality of program

implementation (MENTOR/National Mentoring Partnership, 2003; Weinberger, 2005). A core feature of the program referred to previously is the joint participation of mentors and girls in a series of psychoeducational workshops. Components of the intervention that will undergo piloting include the workshops, program orientation and training procedures, between-session activities to be completed by mentors and youth that are keyed to workshop content, and protocol used for supervision of relationships. Other components of the program (e.g., matching of youth with mentors) will make use of well-established procedures within BBBS agencies and thus will not require piloting. This latter consideration highlights the value that can accrue from utilizing existing program models as a foundation for intervention development efforts (DuBois & Silverthorn, 2005a), especially for programs such as BBBS that have been indicated to have a positive impact on youth outcomes within a controlled evaluation trial (Grossman & Tierney, 1998). Such possibilities illustrate one avenue through which there can be a positive feedback loop between intervention research and preintervention research.

For piloting efforts to be of maximal usefulness in guiding the refinement of different components of programs, it is essential that all aspects of the implementation process be evaluated using appropriate measures, which include the types of assessments of program fidelity and dosage that are described in the following section. Indeed, the piloting process provides a valuable opportunity to refine such measures on the basis of both preliminary examination of their psychometric properties and more practical considerations, such as their acceptability to respondents and the costs involved in administration and implementation.

## PREVENTIVE INTERVENTION RESEARCH

The intervention phase of the research cycle, as applied to youth mentoring, is concerned primarily with establishing the impact of mentoring interventions on participating youth. Evaluations need to be informed by a careful assessment of both intervention strength and fidelity. Evaluations, furthermore, ideally should incorporate data on program costs and on potential cost offsets associated with reduced use of social, health, and criminal justice services by mentored youth, as well as by mentors, and should make use of these to conduct cost-effectiveness and cost-benefit analyses. The following sections address each of these concerns: evaluation of program impact, assessment of program fidelity and dosage, and cost-effectiveness/cost-benefit analysis.

### *Evaluation of Program Impact*

Our focus with regard to evaluating program impact is on issues involved in conducting randomized experimental trials to evaluate the efficacy or effectiveness of mentoring interventions (the interested reader is referred to Grossman, 2005, for a discussion of alternatives to random assignment in the evaluation of mentoring programs). The unique benefits of randomized controlled designs for evaluating program efficacy and effectiveness are emphasized in frameworks for prevention research (Flay, 1986; IOM, 1994; NAMHC, 2001) and in the *Standards of Evidence* adopted in 2004 by the Society for Prevention Research (n.d.).

Several aspects of existing research on mentoring programs, furthermore, point toward a need for experimental designs to advance understanding of their impacts on

youth (DuBois, Holloway, Valentine, & Cooper, 2002; Jekielek et al., 2002; Rhodes, 2002). These include evidence that the salutary effects of mentoring programs as currently constituted, although not necessarily lacking in public health or policy significance, are relatively small in magnitude according to conventional metrics (DuBois et al., 2002; Rhodes, 2002). It is thus difficult to make the case that the likely program impacts are so substantial that the research method used to estimate impacts does not really matter; rather, the sensitivity afforded by a controlled experimental design may be essential for detecting and accurately gauging impacts. Further noteworthy trends include (1) evidence that effects tend to vary by subgroups, including stronger effects for youth facing environmental risk and fewer benefits, and even adverse effects, for youth with substantial personal problems, who perhaps are in need for more intensive services and support than are made available through mentoring programs, at least as currently configured; (2) findings indicating that a range of program design features and implementation practices are correlated with mentoring interventions that have more substantial positive impact on youth; and (3) a lack of evidence on longer-term effects of mentoring, with most research restricted to the period of program participation. The implications of these trends for use of experimental designs in intervention research on mentoring are addressed later.

First, however, challenging research design issues that need to be confronted on a case-by-case basis when using random assignment within evaluations of mentoring programs merit consideration. Among the most salient are the need for excess demand for available program slots and the implications of this need for program operations and the proper point in the program application and selection process at which to place the random assignment lottery. As illustrated later, these issues are likely to be resolved most effectively through collaborative decision making and negotiation between researchers and those in community settings responsible for implementing the program.

Excess demand for program slots is essential for random assignment studies. Program operators must either already have more appropriate applicants than can be served or be able to generate additional applicants through more outreach to produce this surplus. Making this happen can require added resources for recruitment (which must be compensated), because programs often devote just enough effort to recruitment to fill available slots. Furthermore, program operators often find it difficult to tell appropriate applicants, "No, we cannot serve you," preferring to put excess applicants on waiting lists or admitting them for a later period. Referral to waiting lists, if done on a randomized basis, can provide the basis for implementing an experimental evaluation (e.g., Grossman & Tierney, 1998). It has the significant limitation, however, of restricting the opportunities for assessment of longer-term impact because those in the control group (i.e., waiting list) will become eligible for program services within a defined period (e.g., 18 months in the Public/Private Ventures evaluation of the Big Brothers Big Sisters program).

The placement of random assignment in the application/selection process relates to three issues: the administrative burden on programs arising from the study, the composition of the research sample, and the likely difference in services between the program group (given access to the services under study) and the control group (who will look elsewhere for services). Before random assignment, program staff members have to process applications for more people than they will actually serve because they need to assure an applicant pool sufficient to allow for creation of the control group through the lottery. For example, if the random assignment ratio is 50% for the program group and 50% for the control group, then at least twice as many applicants as can be served must

reach the point of random assignment. As the point of random assignment is made later in the application process, program staff will need to process this larger-than-normal group of applicants through more stages of the process. This necessity increases the administrative burden on program staff and makes it more difficult to tell members of the control group (who invest time and energy at each step of the process) that they will not be served.

Placement of the lottery also affects who is in the research sample. In a typical random assignment study, random assignment occurs after staff have recruited a pool of interested applicants, made their usual assessment of the appropriateness of the program services for individuals, and confirmed their interest in participating in the program. Thus, the study will generate estimates of the program's impact for the type of clients usually served. However, at times, random assignment occurs earlier in the application process, for example, because the study is designed to estimate program impacts for a group that includes those less motivated (to follow through on all the steps to apply) or less screened to see whether they meet the usual program requirements than is ordinarily the case. This question could be important if funders wish to understand whether a mentoring program could be effective if it recruited and served a harder-to-serve clientele. In this case, once applicants are randomly assigned to the program group, program staff would make special efforts to encourage participation, especially among the harder-to-serve subgroup. In an extreme case, random assignment could take place even earlier, in order to see whether the program could recruit youth who would not ordinarily be referred or apply on their own initiative for the program and, if they were recruited, whether the program generated positive impacts. At the other end of the spectrum, if random assignment is delayed until late in the application process, only those applicants motivated enough to persist through all the steps will be in the research sample.

Scheduling of the lottery also affects the likely difference in services between the program and control group, and this difference is what "generates" any program impacts: The greater the difference in services, the more likely impacts will be found. In a random assignment impact study, a positive program impact is produced *if* the service being tested is effective *and* the program group receives more of it than the control group does. So the likelihood of finding a statistically significant program impact if the service is effective depends on the "service contrast" between the program and control group. Even if a service is amazingly effective, if the program group and control group participate in similar amounts in similar services, the impact estimate is likely to be zero.[1]

At times, researchers who design random assignment studies and program staff think about the question of where to position the point of random assignment from different perspectives. Program staff may want the program group to have a very high participation rate in the service being tested, seeing this rate as an important part of designing a fair test, so they push for late random assignment to increase the chances people in the program group will actually participate. But designers of the study may be

[1] In addition to placement of the lottery, the size of the service contrast can be influenced by youth in the control group receiving mentoring services. This may occur because youth in the control group enroll in an alternative mentoring program and, in some cases, because they inadvertently are permitted to participate in the program being evaluated (i.e., the integrity of the random-assignment design is compromised). These considerations underscore the importance of assessing the mentoring services received by youth not only in the intervention group, but also in the control group. The service contrast derived from these data may, in fact, be a better predictor of outcomes in some instances than the fidelity of implementation. In places where it is hard to organize and do things, fidelity may not look great, but the difference between the mentoring services that the program and control group receive could be large because the controls have so little access to other sources of mentoring.

focused on the need for a strong service contrast and know that the participation rates of both the program and control group are likely to increase as the point of random assignment is moved to occur later in the application process. (With late random assignment, the sample will be made up of motivated applicants, and those who end up in the control group are more likely to seek alternative services.) Thus, those designing the research will be trying to pick the best way to balance this tradeoff in light of the research question to be addressed.

In sum, although the perceived need for random assignment does build support at the top of organizations for experimental research, many tough questions remain to be addressed and random assignment will not turn out to be feasible everywhere. To date, random assignment studies of mentoring typically have not described their rationale for where to position the point of random assignment (for an exception, see Grossman & Tierney, 1998); nor have syntheses of this research couched interpretation of findings within the context of a consideration of decisions made in this regard.

As mentioned earlier, the existing research suggests that mentoring programs have different effects for young people that depend on the environmental risk factors they face and the extent of their own personal problems. These trends suggest that it is important for evaluations of mentoring programs to include a diverse sample of youth, one that encompasses the full range of different levels and configurations of risk that are typical of the youth who are intended to be served by the program. This emphasis can be contrary to the instincts of program operators who—when faced with an evaluation—may want the research sample to include youngsters who are likely to show good outcomes during and after their participation in the program. Although such a sample would show the program in a good light if the focus were on the outcomes of youngsters served (e.g., the percentage who attend school regularly), it could cause problems when the research question was the program's impact—the difference it made in outcomes. Therefore, researchers need to help program operators see this counterintuitive lesson from the previous research.

As noted, existing research also suggests different "best practices" that may enhance the impact of mentoring programs for youth. These practices have been identified almost exclusively through the nonexperimental analysis of findings across different evaluations (see, in particular, DuBois, Holloway, et al., 2002). Direct experimental studies of variations in program design and practice clearly would be preferable and could significantly advance knowledge in this area. Experimental tests could be structured, for example, to compare different strategies for recruiting and matching mentors and youngsters or different durations of relationships. Applicants could be randomly assigned to two or even three different program groups (which vary on key aspects of program design) or to a control group. Conclusions about the relative effectiveness of different programmatic approaches then could be drawn more reliably because the youngsters served in each would be truly comparable (Grossman, 2005). It is important to note, however, that these differential impact tests do require substantially larger samples than simple two-way comparisons.

Extending follow-up past the point of participation in mentoring, even into adulthood, will provide valuable new insights into the effects of programs. This type of recommendation is made frequently when considering potential areas of improvement for evaluations of youth programs. But with mentoring, there is reason to believe the programs could make a real long-term difference on the basis of the short-term findings. A sound basis thus exists to argue for the investment in evaluations that afford the opportunity to evaluate long-term effects of mentoring.

Finally, as pivotal as random assignment is widely held to be for establishing program efficacy or effectiveness, it is nonetheless equally important for those conducting research on mentoring interventions to keep in mind that random assignment is best regarded as a necessary, but not sufficient, condition for supporting claims of program impact. The *Standards of Evidence* referred to previously include numerous additional criteria that must be met to support conclusions of program efficacy or effectiveness (SPR, n.d.). These include, but are not limited to, such considerations as utilizing multiple sources of data when "demand characteristics" are plausible for measures, ensuring that analyses are carried out at the same level as randomization, correcting for increases in type I error rate adjustment when analyzing multiple outcomes, conducting analyses that take into account potential bias caused by differential attrition, demonstrating practical significance of program effects in terms of public health impact, and, for outcomes that may decay over time (and this category arguably applies to nearly all outcomes of concern to mentoring programs), establishing maintenance of significant effects at one or more long-term follow-up assessments. It does not appear that any of the existing random assignment studies of mentoring programs have met the full complement of these criteria. It is even more certain that none has met the further criterion that consistent findings be reported across at least two different high-quality studies that meet all other criteria (SPR, n.d.). Clearly, there is much work ahead for those who seek to establish claims of efficacy or effectiveness for mentoring programs that will be embraced by the larger scientific prevention community.

### *Assessing Intervention Strength and Fidelity*

All evaluations of youth mentoring programs need to be informed by careful assessments of intervention strength and fidelity. Intervention strength, also known as *treatment or intervention exposure* (Rohrbach, Graham, & Hansen, 1993), is defined as the "dose, duration, specificity, and intensity" of a given intervention (Summerfeldt, 2003, p. 56), which may vary across different participants in an intervention. In contrast, *intervention fidelity* refers to the extent to which an intervention is actually implemented as planned (Gottfredson, Gottfredson, & Skroban, 1998; Summerfeldt, 2003). Also described as "program adherence" (Center for Substance Abuse Prevention [CSAP], 2002) or "program integrity" (Orwin, 2000), *fidelity* is sometimes used as an overarching term to refer to both strength and fidelity as defined here (CSAP, 2002). In this article, as in most other discussions of the topic (Sechrest, Phillips, Redner, & Yeaton, 1979), however, each construct is assumed to be a distinct component of program implementation.

Although there are many reasons why a social or preventive intervention is effective, one critical factor is the quality of a program's implementation (Tebes, Kaufman, & Connell, 2003). This finding has been observed repeatedly in a variety of studies (Blakely et al., 1987; CSAP, 2002), including in the youth mentoring literature (DuBois, Holloway et al., 2002). Despite these findings, attention to intervention strength and fidelity is frequently neglected when designing and evaluating social and preventive interventions (Summerfeldt, 2003). Under the best of circumstances, evaluators of prevention programs complete several steps to assess intervention strength and fidelity. These include: conducting a component analysis, establishing implementation standards, measuring intervention strength and fidelity, and examining reasons for failing to implement the program as designed (Gottfredson et al., 1998; Scheirer, 1994; Tebes et al., 2003). The sections that follow briefly discuss each of these processes with particular attention to youth mentoring programs.

*Conducting a Component Analysis.* The core components of a program are drawn directly from the overarching program theory and the program's logic model. What experiences, behaviors, and events are expected to result from specific program activities, and how are they related to short-, intermediate-, and long-term outcomes? One may conduct a core component analysis through the use of primary or secondary analysis of program materials, observations of the program as it is implemented, or interviews with key program stakeholders about the program (McGrew, Bond, Dietzen, & Salyers, 1994). What guides such an analysis, however, is the program logic model. Thus far, there have been few component analyses conducted in the youth mentoring field, perhaps because few mentoring programs have adequately specified a program logic model. The absence of clearly specified logic models that describe how, and under what conditions, the program and mentor–protégé relationships established within the program are expected to lead to positive change among youth is undoubtedly related to the lack of formal theory underlying most efforts to develop mentoring programs (DuBois & Karcher, 2005). Theories that have begun to appear and be tested in the mentoring research literature (e.g., DuBois, Neville, et al., 2002; Rhodes, 2002, 2005) could be used to inform the development and evaluation of mentoring programs, but, to date, there is little evidence of this process.

*Establishing Implementation Standards.* Investigators intent on measuring intervention strength and fidelity usually establish standards for each component of their prevention program model (Gottfredson et al., 1998; Orwin, 2000). Standards represent levels at which the investigator believes that the intervention will have an impact on its intended targets,, such as youth. For example, an investigator may identify minimal levels of dosage a youth should receive for the program to be effective and/or the duration of time specific intervention activities should be delivered to have some likelihood of being effective. The investigator may also describe the specific activities that should happen at specific times and at a specific intensity in order for the intervention to be effective. To the extent possible, these standards are derived from previous related research, the program developer's experience in implementing similar programs, and the opinions of key stakeholders involved with the program, such as the youth who are intended targets of the intervention, their parents, mentors, or other key individuals (such as school personnel or community members) whose support for the program may be crucial to its success.

For the most part, specification of implementation standards for youth mentoring programs that have been evaluated in the literature has been centered on program criteria for minimally acceptable levels of the frequency of mentor–youth contact and the duration of relationships (Rhodes, 2002). Whether or not program standards are met in these areas has been demonstrated to have important implications for youth outcomes (e.g., DuBois, Neville, et al., 2002; Grossman & Rhodes, 2002). Much less attention has been given to specification of implementation standards that pertain to other potentially important dimensions of mentoring relationships, such as the content of mentor–youth activities and discussions together and the mentor's use of different types of strategies for promoting youth outcomes (e.g., goal setting). The practice and research literatures also highlight program implementation processes in several areas that may be important for both (1) establishing and supporting effective relationships, such as mentor recruitment and screening, matching of youth and mentors, training, and supervision (MENTOR/National Mentoring Partnership, 2003; Rhodes, 2002), and (2) ensuring appropriate linkages and coordination between mentoring and other components or services offered within multicomponent interventions (Kuperminc et al., 2005). With notable exceptions

(e.g., Taylor, LoSciuto, Fox, & Hilbert, 1999), implementation standards for these latter types of relationship and program elements have not been articulated by either program developers or evaluators.

*Measuring Intervention Strength and Fidelity.* Intervention strength and fidelity may be measured by using a variety of data sources (Gottfredson et al., 1998; McGrew et al., 1994), such as reviews of contact logs, interviews with program implementers and recipients, observations and ratings of program activities, and examination of archival program data, including budgets (McGrew et al., 1994; Orwin, 2000). In the case of youth mentoring programs, assessments of intervention strength and fidelity, for example, may include the number of meetings between a mentor and protégé that take place within a given period. This type of assessment allows for calculation of a ratio, or score, that depicts the number of meetings held as a proportion of those prescribed. Similar scores are possible, in principle, for any of the other relationship and program factors described previously for which implementation standards have been established (for program factors, associations with variations in outcome may be most feasible to examine in the context of evaluations of programs implemented across multiple sites). These scores can then be examined for their relationships to various program outcomes. To the extent that the program's logic model describes the influence of several core components (as is expected in the case of mentoring programs), such scores also can be examined in combination for their relationship to outcome (Gottfredson et al., 1998). Examination of these types of associations provides a valuable mechanism for generating hypotheses about the relative and collective importance of different intervention components in producing desired outcomes.

Several studies in the youth mentoring field have used fidelity instruments to assess program implementation and examined associations with youth outcomes. For the most part, these studies have focused on variability in relationship factors. In their meta-analysis, DuBois, Holloway, et al. (2002) synthesized findings from nine of these types of investigations and found that, on average, youth who experienced mentoring relationships of greater intensity or quality in programs scored between one quarter and one third of a standard deviation higher in a favorable direction on outcome measures. In part because of the rarity of multisite evaluations in the mentoring literature, there has been comparatively little corresponding examination of how fidelity of implementation in program-level factors (e.g., training) relates to youth outcomes.

*Examining Reasons for Failing to Implement the Program as Designed.* Investigators who develop a specified program theory and logic model give themselves the best opportunity to implement a program with success and know *why* it worked. However, they must also accept the possibility that they may be wrong. When a program is not implemented as designed, assessing the veracity of the program theory or logic model becomes virtually impossible. That is why careful implementation of prevention programs and the measurement of intervention strength and fidelity are essential. There are an infinite number of reasons for failing to implement a program as designed. One of the most common, however, is that the program design was too complex for successful implementation, or relatedly, insufficient training, supervision, and monitoring were provided to the program implementers (Tebes et al., 2003). In relation to youth mentoring programs, it is widely accepted that orientation and training for mentors, as well as ongoing supervision and support, are critical to a program's success (Sipe, 1996). Equally important, however, may be training and technical assistance provided to administrators and staff responsible for imple-

menting mentoring programs. This need may be heightened when mentoring is made available within a broader, multifaceted intervention, and there are demands to integrate and coordinate mentoring with other programs and services (Kuperminc et al., 2005). Both quantitative and qualitative data that are gathered on the fidelity of the implementation process may offer important insights about factors responsible for failures in implementation for mentoring programs. To date, however, these have received limited attention in the literature (for a noteworthy exception, see Hamilton & Hamilton, 1992).

### *Cost-Effectiveness and Cost-Benefit Analyses*

Measuring, understanding, and improving the costs, benefits, cost-effectiveness, and cost-benefit ratio of youth mentoring programs are central to building a strong case for the dissemination of mentoring to wider audiences of professionals and funders (SPR, n.d.). A full treatment of the issues involved in assessing the cost-effectiveness and cost-benefit ratio of youth mentoring programs is beyond the scope of this article (for an in-depth examination, see Yates, 2005). Our focus here is to frame a few of the conceptual issues involved with these types of analyses and to comment briefly on existing efforts to conduct cost-effectiveness and cost-benefit analyses of youth mentoring programs.

There is a view among youth mentoring program propagators that the key advantages of mentoring programs over other programs include (1) the relatively low *cost* of using mentors; (2) the benefits of reduced use of social, health, educational, and criminal justice services by mentored youth; (3) the equivalency, or superiority, of the *effectiveness* of mentors in providing some or all services, relative to professionals; and hence, (4) the *cost-effectiveness* of using mentors as opposed to professional staff for some or all services—and, given the perceived low cost or lack of cost of mentors, (5) the near guarantee that the *benefits* of mentors will exceed the *costs* of mentors, that is, that mentors must be *cost-beneficial*. Each of these perceived advantages can be criticized; because each is measurable, each also can be examined through quantitative assessment. For example, from the perspective of community representatives and funders, paid and volunteer mentors may be thought of as costing less than paid staff—or even nothing—for a program. However, paid staff of some programs spend considerable time recruiting, conducting background checks for, matching, training, assigning, and monitoring mentors and the youth with whom they work. Similarly, mentors are often thought of as providing services that either are unique or may substitute for some services that professionals might otherwise have to provide, but the effectiveness of mentors' services can be challenged (e.g., Blechman & Bopp, 2005).

The only viable strategy for resolving questions about the costs and effectiveness (or benefits) of mentoring programs is to measure each. To do so, it is important to consider costs, effects, and benefits comprehensively from the perspectives of different interest groups, including program managers, community members, mentors, and youth (Yates, 2005). Aside from being important from the standpoint of conducting a thorough cost-effectiveness or cost-benefit analysis, consideration of each of these groups may yield valuable insights that can inform the further development of mentoring programs. For instance, why do mentors work so hard for so little? One answer is that they actually work hard for a great deal. Mentors may receive benefits that typically exceed the costs of time not spent working or being with family and friends, transportation to and from protégés' locations, and incidental expenses, such as supplies and admission to museums or parks. Mentors can be thought of as a particular type of

volunteer. It has been recognized for some time that the benefits received by volunteers (e.g., training, education, personal insight, gratification from helping others) can equal or exceed the costs borne by volunteers (e.g., earnings forgone because of time spent volunteering rather than working, transportation expenses paid by volunteers) (see, e.g., Yates, 1980). In traditional exchange theory, benefits can exceed costs in a social interaction for only a subset of participants, with the other participants experiencing more costs than benefits. When benefits and costs are measured from the perspective of each party in the interaction, however, costs can diminish and benefits can increase so that all parties experience benefits that exceed their costs.

To date, there has been very limited attention to cost-effectiveness and cost-benefit analyses in evaluations of youth mentoring programs (for a review, see Yates, 2005). Methodological limitations of existing efforts include both a failure to consider costs and effects/benefits from the multiple perspectives referred to earlier (e.g., both mentor and youth; Blechman, Maurice, Buecker, & Helberg, 2000), as well as the inherent difficulties associated with attempting to estimate cost-effectiveness and cost-benefit ratios from evaluations that were not necessarily originally designed with this aim in mind (Aos, Lieb, Mayfield, Miller, & Pennucci, 2004). In view of these considerations and the small size of the available literature, Yates (2005) concluded that reliable data on cost-effectiveness and cost-benefit ratio for mentoring programs are not currently available.

## PREVENTIVE SERVICE SYSTEMS RESEARCH

As noted in Figure 1, dissemination and implementation studies are a major focus of preventive service systems research on programs that have been found to be efficacious and effective in intervention. Dissemination studies typically are based on data collected from representative samples of potential host organizations and agencies and may utilize both nonexperimental and experimental designs (Oldenburg & Parcel, 2002). The youth mentoring literature currently lacks studies of dissemination and the other types of related research referred to in Figure 1 as being encompassed under the category of preventive service systems research. As effective mentoring programs are identified through research at the intervention stage, however, these types of investigations have the potential to serve a variety of important purposes. Nonexperimental studies could collect quantitative and qualitative data to investigate predictors of differences in the adoption, implementation, and institutionalization of effective programs by host organizations and systems, as well as to increase understanding of specific barriers and facilitators of effective dissemination. Experimentally designed studies (as well as those utilizing well-controlled quasi-experimental designs), furthermore, could be used rigorously to evaluate the effectiveness of differing approaches to the dissemination of effective mentoring programs to youth-serving organizations. Dissemination studies also may be linked in useful ways with outcome-based evaluation research, such as large-scale effectiveness or demonstration projects. When combined with outcome data, for example, dissemination data may be helpful in deriving population-based estimates of the likely impact of selected programs or practices. These types of investigations, furthermore, could permit the examination of how organizational processes that occur at multiple levels influence the ultimate effects of mentoring programs and practices on youth outcomes. In this way, youth mentoring research may serve as an exemplar for studies in the emerging field of community science in which multilevel processes are examined in individual, interpersonal, and community contexts (Tebes, 2005).

## CONCLUSIONS AND FUTURE DIRECTIONS

### *Conclusions*

Research on youth mentoring, despite many noteworthy accomplishments to date, remains in an early stage of development (DuBois & Karcher, 2005). Our examination of the current state of research methodology in the field within the framework of the different recommended phases of preventive intervention research highlights several important issues and approaches that merit greater attention at the preintervention, intervention, and systems phases of investigation. Future work that addresses these gaps and limitations holds promise of significantly advancing the field's knowledge base and, ultimately, the capacity of mentoring interventions to have a substantial and lasting impact on youth outcomes. Even more noteworthy to us, however, is the relative absence of programmatic research that reflects an integrated progression across the different phases of research in development, evaluation, and dissemination. Our recommendations for future research thus focus not only on methodological issues specific to each phase, but also on the need for increased linkage and coordination of work across phases.

### *Recommendations for Research*

*Preintervention Research.* Basic research on youth mentoring relationships will benefit from utilizing, whenever possible, (1) large, representative samples so as to maximize sensitivity to relationship dynamics and generalizability of findings; (2) longitudinal designs that include both numbers and time frames of assessments that are most suitable to addressing questions of interest; (3) multiple sources and methods for assessing mentoring relationships to triangulate more accurately their most influential characteristics and processes; and (4) sophisticated multivariate data analytic procedures, particularly those that are appropriate for model testing and examination of phenomena that occur at the dyadic level in relationships. Basic research needs to be complemented by initial research on the development of mentoring interventions focused on obtaining and utilizing stakeholder input, piloting intervention strategies, and development of psychometrically sound evaluation protocols.

*Intervention Research.* Intervention research should give priority to conducting (1) experimental trials of efficacy and effectiveness that satisfy the most methodologically stringent criteria of acceptability (SPR, n.d.) and that are structured, when possible, to allow experimental tests of the impacts associated with specific practices and procedures within programs; (2) theoretically informed assessments of intervention strength and fidelity and their implications for program effectiveness that incorporate attention to each of the specific steps discussed earlier (i.e., conducting a component analysis, establishing implementation standards, measuring intervention strength and fidelity, and examining reasons for implementation failures); and (3) proactively designed cost-benefit and cost-effectiveness analyses that incorporate assessments of costs and benefits for multiple groups, including both youth and mentors.

*Preventive Service Systems Research.* Factors that affect the adoption, implementation, and institutionalization of effective mentoring programs within larger systems should be

investigated by using both (1) nonexperimental methods to study rates and predictors of naturally occurring variations in dissemination and (2) experimental designs that allow innovations in approaches to dissemination to be rigorously tested and examined for potential population-level impacts.

*Linking Phases of Research.* The highest priority should be given to conducting programmatic research that addresses a critical need for stronger linkages among the different phases of investigation of youth mentoring interventions. To date, many of the most widely disseminated programs lack a well-articulated foundation in the field's empirical knowledge base. This may, in part, reflect the fact that research typically is university based, whereas mentoring programs for youth most often have been developed by community-based, grass-roots organizations, the efforts of many of which predate some of the most noteworthy empirical advances in the field. Desirable linkages are also lacking between the phases focused on evaluation and on dissemination. As a result, some of the programs with the most promising evaluation data have had only limited dissemination. Conversely, several programs have been the focus of noteworthy dissemination efforts in the absence of well-designed evaluation research. To address these concerns, future research on youth mentoring will need to be supported by strong networks of communication that link the efforts of different investigators not only with each other, but also with practitioners and policy makers.

## REFERENCES

AOL Time Warner Foundation. (2002). Mentoring in America 2002. Retrieved July 20, 2004, from http://mentoring.web.aol.com/common/one_report/national_poll_report_final.pdf

Aos, S., Lieb, R., Mayfield, J., Miller, M., & Pennucci, A. (2004). Benefits and costs of prevention and early intervention programs for youth. Report to the Washington State Legislature. Olympia: Washington State Institute for Public Policy. Retrieved July 18, 2004, from http://www.wsipp.wa.gov/

Baker, D.B., & Maguire, C.P. (2005). Mentoring in historical perspective. In D.L. DuBois & M.J. Karcher (Eds.), Handbook of youth mentoring (pp. 14–29). Thousand Oaks, CA: Sage.

Bartholomew, L.K., Parcel, G.S., Kok, G., & Gottlieb, N.H. (2001). Intervention mapping: Designing theory- and evidence-based health promotion programs. London: Mayfield.

Bearman, P.S., & Jones J., & Udry, J.R. (2003). The National Longitudinal Study of Adolescent Health: Research design. Retrieved June 20, 2003, from http://www.cpc.unc.edu/addhealth

Blakely, C.H., Mayer, J.P., Gottschalk, R.G., Schmitt, N., Davidson, W.S., Rollman, D.B., et al. (1987). The fidelity-adaptation debate: Implications for the implementation of public sector social programs. American Journal of Community Psychology, 15, 253–269.

Blechman, E., Maurice, A., Buecker, B., & Helberg, C. (2000). Can mentoring or skill training reduce recidivism? Observational study with propensity analysis. Prevention Science, 1, 139–155.

Blechman, E.A., & Bopp, J.M. (2005). Juvenile offenders. In D.L. DuBois & M.J. Karcher (Eds.), Handbook of youth mentoring (pp. 454–466). Thousand Oaks, CA: Sage.

Center for Substance Abuse Prevention. (2002). Finding the balance: Program fidelity and adaptation in substance abuse prevention. DHHS 2002 Conference Edition. Rockville, MD: Author.

Coyne, S.M., Duffy, J.L., & Wandersman, A. (2005). Mentoring for results: Accountability at individual, program, community, and policy levels. In D.L. DuBois & M.J. Karcher (Eds.), Handbook of youth mentoring (pp. 546–560). Thousand Oaks, CA: Sage.

Dalton, J.H., Elias, M.J., & Wandersman, A. (2001). Community psychology: Linking individuals and community. Belmont, CA: Wadsworth/Thompson Learning.

Darling, N., Hamilton, S., Toyokawa, T., & Matsuda, S. (2002). Naturally occurring mentoring in Japan and the United States: Roles and correlates. American Journal of Community Psychology, 30, 245–270.

DuBois, D.L., Holloway, B.E., Valentine, J.C., & Cooper, H. (2002). Effectiveness of mentoring programs for youth: A meta-analytic review. American Journal of Community Psychology, 30, 157–197.

DuBois, D.L., & Karcher, M.J. (2005). Youth mentoring: Theory, research, and practice. In D.L. DuBois & M.J. Karcher (Eds.), Handbook of youth mentoring (pp. 2–12). Thousand Oaks, CA: Sage.

DuBois, D.L., Neville, H.A., Parra, G.R., & Pugh-Lilly, A.O. (2002). Testing a new model of mentoring. New Directions for Youth Development, 93, 21–57.

DuBois, D.L., & Silverthorn, N. (2005a). Research methodology. In D.L. DuBois & M.J. Karcher (Eds.), Handbook of youth mentoring (pp. 44–64). Thousand Oaks, CA: Sage.

DuBois, D.L., & Silverthorn, N. (2005b). Characteristics of natural mentoring relationships and adolescent adjustment: Evidence from a national study. Journal of Primary Prevention, 26, 69–92.

DuBois, D.L., & Silverthorn, N. (2005c). Natural mentoring relationships and adolescent health: Evidence from a national study. American Journal of Public Health, 95, 518–524.

Flay, B.R. (1986). Efficacy and effectiveness trials (and other phases of research) in the development of health promotion programs. Preventive Medicine, 15, 451–474.

Gottfredson, D.C., Gottfredson, G.D., & Skroban, S. (1998). Can prevention work where it is needed most? Evaluation Review, 22, 315–340.

Green, L.W., & Kreuter, M.W. (1999). Health promotion planning: An educational and environmental approach (2nd ed.). Mountain View, CA: Mayfield.

Grossman, J.B. (2005). Evaluating mentor programs. In D.L. DuBois & M.J. Karcher (Eds.), Handbook of youth mentoring. (pp. 251–265). Thousand Oaks, CA: Sage.

Grossman, J.B., & Rhodes, J.E. (2002). The test of time: Predictors and effects of duration in youth mentoring programs. American Journal of Community Psychology 30, 199–206.

Grossman, J.B., & Tierney, J.P. (1998). Does mentoring work? An impact study of the Big Brothers/Big Sisters program. program. Evaluation Review, 22, 403–426.

Hamilton, M.A., & Hamilton, S.F. (2005). Work and service-learning. In D.L. DuBois & M.J. Karcher (Eds.), Handbook of youth mentoring (pp. 14–29). Thousand Oaks, CA: Sage.

Hamilton, S.F., & Hamilton, M.A. (1992). Mentoring programs: Promise and paradox. Phi Delta Kappan, 73, 546–550.

Herrera, C.L., Vang, Z., & Gale, L.Y. (2002). Group mentoring: A study of mentoring groups in three programs. Philadelphia: Public/Private Ventures.

Hirsch, B.J. (2005). A place called home: Afterschool programs for urban youth. Washington, DC: American Psychological Association and New York: Teachers College Press.

Hirsch, B.J., & Wong, V. (2005). After-school programs. In D.L. DuBois & M.J. Karcher (Eds.), Handbook of youth mentoring (pp. 14–29). Thousand Oaks, CA: Sage.

Institute of Medicine. (1994). Reducing risks for mental disorders: Frontiers for preventive intervention research. Washington, DC: National Academy Press.

Jekielek, S., Moore, K.A., & Hair, E.C. (2002). Mentoring programs and youth development. Washington, DC: Child Trends.

Karcher, M.J., Nakkula, M.J., & Harris, J.T. (2005). Developmental mentoring match characteristics: Correspondence between mentors' and mentees' assessments of relationship quality. Journal of Primary Prevention, 26, 93–110.

Keller, T.E. (2005). The stages and development of mentoring relationships. In D.L. DuBois & M.J. Karcher (Eds.), Handbook of youth mentoring (pp. 82–99). Thousand Oaks, CA: Sage.

Kuperminc, G.P., Emshoff, J.G., Reiner, M.M., Secrest, L.A., Niolon, P.H., & Foster, J.D. (2005). Integration of mentoring with other programs and services. In D.L. DuBois & M.J. Karcher (Eds.), Handbook of youth mentoring (pp. 314–334). Thousand Oaks, CA: Sage.

Langhout, R.D., Rhodes, J.E., & Osborne, L. (2004). An exploratory study of youth mentoring in an urban context: Adolescents' perceptions of relationship styles. Journal of Youth and Adolescence, 33, 293–306.

Liang, B., Tracy, A.J., Taylor, C.A., & Williams, L.M. (2002). Mentoring college-age women: A relational approach. American Journal of Community Psychology, 30, 271–288.

McGrew, J.H., Bond, G.R., Dietzen, L., & Salyers, M. (1994). Measuring the fidelity of implementation of a mental health program model. Journal of Consulting and Clinical Psychology, 62, 670–678.

McLearn, K.T., Colasanto, D., & Schoen, C. (1998). Mentoring makes a difference: Findings from the Commonwealth Fund 1998 Survey of Adults Mentoring Young People. New York: Commonwealth Fund.

MENTOR/National Mentoring Partnership. (2003). Elements of effective practice (2nd ed.). Alexandria, VA: Author.

Nakkula, M.J., & Harris, J.T. (2005). Assessment of mentoring relationships. In D.L. DuBois & M.J. Karcher (Eds.), Handbook of youth mentoring (pp. 100–117). Thousand Oaks, CA: Sage.

National Advisory Mental Health Council Workgroup on Mental Disorders Prevention Research. (2001). Priorities for prevention research at NIMH. Prevention & Treatment, 4, article 17. Retrieved July 14, 2004, from http://www.journals.apa.org/prevention/volume4/pre0040017a.html

National Institute of Mental Health. (1998). Priorities for prevention research at NIMH: A report by the National Advisory Mental Health Council (NIH Publication No. 98-2079). Bethesda, MD: Author.

Newman, S., Morris, G.A., & Streetman, H. (1999). Elder-child interaction analysis: An observation instrument for classrooms involving older adults as mentors, tutors or resource persons. Child & Youth Services, 20, 129–145.

Oldenburg, B., & Parcel, G.S. (2002). Diffusion of innovations. In K. Glanz, B.K. Rimer, & F.M. Lewis (Eds.), Health behavior and health education: Theory, research, and practice (3rd ed., pp. 312–334). San Francisco: Jossey-Bass.

Orwin, R.G. (2000). Assessing program fidelity in substance abuse health services research. Addiction, 95(Supp. 3), S309–S327.

Parra, G.R., DuBois, D.L., Neville, H.A., Pugh-Lilly, A.O., & Povinelli, N. (2002). Mentoring relationships for youth: Investigation of a process-oriented model. Journal of Community Psychology, 30, 367–388.

Portwood, S.G., & Ayers, P.M. (2005). Schools. In D.L. DuBois & M.J. Karcher (Eds.), Handbook of youth mentoring (pp. 336–347). Thousand Oaks, CA: Sage.

Rhodes, J.E. (2002). Stand by me: The risks and rewards of mentoring today's youth. Cambridge, MA: Harvard University Press.

Rhodes, J.E. (2005). A model of youth mentoring. In D.L. DuBois & M.J. Karcher (Eds.), Handbook of youth mentoring (pp. 30–43). Thousand Oaks, CA: Sage.

Rhodes, J.E., & DuBois, D. (2004). National research agenda for youth mentoring. Alexandria, VA: MENTOR/National Mentoring Partnership. Retrieved July 14, 2004, from http://www.mentoring.org/research_corner/researchagenda.pdf

Rhodes, J.E., Grossman, J.B., & Resch, N.R. (2000). Agents of change: Pathways through which mentoring relationships influence adolescents' academic adjustment. Child Development, 71, 1662–1671.

Rohrbach, L.A., Graham, J.W., & Hansen, W.B. (1993). Diffusion of a school-based substance abuse prevention program: Predictors of program implementation. Preventive Medicine, 22, 237–260.

Scheirer, M.A. (1994). Designing and using process evaluation. In J.S. Wholey, H.P. Hatry, & K.E. Newcomer (Eds.), Handbook of practical program evaluation (pp. 40–68). San Francisco: Jossey-Bass.

Sechrest, L., Phillips, M.A., Redner, R., & Yeaton, W. (1979). Some neglected problems in evaluation research: Strength and integrity of treatments. Evaluation Studies Review Annual, 4, 14–35.

Sipe, C.L. (1996). Mentoring: A synthesis of Public/Private Venture's research: 1988–1995. Philadelphia: Public/Private Ventures.

Sipe, C.L. (2005). Toward a typology of mentoring. In D.L. DuBois & M.J. Karcher (Eds.), Handbook of youth mentoring (pp. 65–79). Thousand Oaks, CA: Sage.

Society for Prevention Research. (n.d.). Standards of evidence: Criteria for efficacy, effectiveness, and dissemination. Retrieved July 14, 2004, from http://www.preventionresearch.org/standardsofevidencebook.pdf

Spencer, R., & Rhodes, J.E. (2005). A counseling and psychotherapy perspective on mentoring relationships. In D.L. DuBois & M.J. Karcher (Eds.), Handbook of youth mentoring (pp. 118–132). Thousand Oaks, CA: Sage.

Summerfeldt, W.T. (2003). Program strength and fidelity in evaluation. Applied Developmental Science, 7, 55–61.

Taylor A., LoSciuto L., Fox, M., & Hilbert S. (1999). The mentoring factor: An evaluation of Across Ages. In V.S. Kuehne (Ed.), Intergenerational program research: Understanding what we have created (pp. 77–100). Binghamton, NY: Haworth Press.

Tebes, J.K. (2005). Community science, philosophy of science, and the practice of research. American Journal of Community Psychology, 35, 213–230.

Tebes, J.K., Kaufman, J.S., & Connell, C. (2003). The evaluation of prevention and health promotion programs. In T.P. Gullotta & M. Bloom (Eds.), Encyclopedia of primary prevention and health promotion (pp. 46–63). New York: Kluwer/Academic.

Weinberger, S.G. (2005). Developing a mentoring program. In D.L. DuBois & M.J. Karcher (Eds.), Handbook of youth mentoring (pp. 220–234). Thousand Oaks, CA: Sage.

Yates, B.T. (1980). Benefits and costs of community-academia interaction in a paraprofessional training course. Teaching of Psychology, 7, 8–12.

Yates, B.T. (2005). Cost-benefit and cost-effectiveness analyses. In D.L. DuBois & M.J. Karcher (Eds.), Handbook of youth mentoring (pp. 525–545). Thousand Oaks, CA: Sage.