

# **IN GOD WE TRUST, ALL OTHERS BRING DATA: ASSESSING THE STATE OF OUTCOMES MEASUREMENT FOR FAITH-BASED AND COMMUNITY-BASED PROGRAMMING**

Robert L. Fischer

After being set in motion by the Charitable Choice provision of the Personal Responsibility and Work Opportunity Reconciliation Act of 1996, and expanded and institutionalized by executive orders beginning in 2001, the Faith-Based and Community Initiative (FBCI) has sought to encourage greater participation by faith-based and community organizations (FBCOs) in federally funded human service programming. Since 2001, the FBCI has aimed to give these organizations equal opportunity with secular and larger organizations to secure federal funding for the delivery of social services. The expanded role for FBCOs within the domain of federally funded human services has carried with it an increased interest in the capacity and operational effectiveness of these organizations.

During this period, the existing literature on FBCO programming and research on all aspects of faith in the provision of human services has grown dramatically. One portion of the growing evidence base pertains to the effectiveness of the services delivered by FBCOs. Approaching the task of understanding outcomes in the FBCO arena is not without marked challenge. In their book on FBCO welfare-to-work programs, Monsma and Soper (2006) concluded in a chapter titled “The Effectiveness Muddle” that “effectiveness—even when defined in terms of program outcomes—is a complex concept that is often misunderstood. There are methodological problems in operationalizing a study that seeks to measure program outcomes from different types of programs in order to compare their relative effectiveness. There are few existing theories or studies to guide us. Nevertheless, public policy makers constantly seek to make decisions based on the presumed relative effectiveness of different types of human services programs” (p. 37).

Given the relative youth of the FBCI, and the reality that the development of the capacity of FBCO programs will take time, the growth of the research literature should be seen as evolutionary. In the context of the limited state of current research, important questions need to be addressed about how FBCOs conduct their programs and the extent to which the programs achieve the intended outcomes. The overarching interest of this paper is in programs that have been the emphasis of the federal FBCI, both faith-based and secular. However, this review

concentrates nearly exclusively on faith-based programs, as these programs have been the subject of the most focused study. This study seeks to assess the current evidence base by examining the previous relevant reviews, highlighting noteworthy studies in the field, and conducting a synthesis of the available research on FBCO effectiveness. The analysis places a premium on examining the available evidence for the purpose of effectively informing policy and practice as it relates to the FBCI.

### **CONTEXT FOR RESEARCH ON THE EFFECTIVENESS OF FAITH-BASED AND COMMUNITY-BASED SERVICES**

The FBCI is rooted in the notion of drawing on the natural capacities and strengths of FBCOs to deliver effective programming. These organizations, as indigenous entities with staff who often reside in the surrounding neighborhoods, are seen as having invaluable connections and credibility within their communities. FBCOs often have “direct and consistent contact” with the most needy in their communities and their articulated mission serves to dedicate their efforts to serving these needs as best they can (Fink & Branch, 2005, p. 1). Also, FBCOs have established roots and connections both within the geographic area where they are located as well as within the broader faith communities.

FBCOs are recognized as having particular expertise and advantage in some areas. For example, in its narrative on the Compassion Capital Fund, the U.S. Department of Health and Human Services (DHSS) describes FBCOs as being “uniquely situated” to serve “families in poverty, prisoners reentering the community and their families, children of prisoners, homeless families, and at-risk youth” (DHHS, 2002). Beyond simply acknowledging expertise, however, some proponents outside of the FBCI have gone further, arguing that faith-based programs are in fact superior to the conventional alternatives in terms of effectiveness and cost (Cnaan & Boddie, 2006; Smith, Bartkowski, & Grettenberger, 2005). These claims appear to be based on the experiences of single studies or compelling anecdotes, rather than on a systematic review of the evidence. The dialogue about the prospective effectiveness of FBCO programming has had a tendency to outpace the actual data available, or to focus on effects for subgroups of participants (Singer & Friel, 2007; *Wall Street Journal*, 2003). This situation underscores the need for additional research on the outcomes of FBCO-sponsored programs, and specifically in comparison to conventional programming or the absence of programming entirely.

In particular, the central question of the unique effectiveness of faith-based organizations in providing human services is also confounded by the role that religion or faith may play in the lives of individuals. Simply examining the effects of programs that have faith as an element of

their delivery will not shed light on the central question. The issue is far more circumscribed (see Figure A-1 in Appendix A.) All individuals possess personal characteristics that pertain to their level of faith and their engagement in religious practice. The existing research shows that the association between religious engagement and positive behaviors and well-being is fairly clear—the relationship is significant and positive but not conclusively causative (e.g., Johnson, Tompkins, & Webb, 2002). Although religious engagement is associated with indicators of better health and behaviors, it is also plausible that the individuals who choose healthy behaviors also tend to choose religious engagement (hence the multidirectional arrows in Figure A-1). When the notion of an FBCO program is introduced into these relationships, it becomes apparent that the individual's choice about what services to use may reflect something about their faith disposition. Although the FBCO program may be influencing the behaviors and well-being of the individual, these changes may also be related to the underlying characteristics of the individuals served.

The ability to distinguish between the effect of this so-called selection bias and true treatment effect depends heavily on the research methodology that is undertaken (Berk, 1983). In a study that specifically examined the factors associated with individuals volunteering for a faith-based program, the authors suggest that researchers can only effectively address this issue through more rigorous study designs involving a sound comparison between served and nonserved groups (Camp, Klein-Saffran, Kwon, Daggett, & Joseph, 2006). In addition, beyond the characteristics of the organization there are also the characteristics of the individual program. For example, an FBCO may deliver different programs that possess different levels of faith intensity, and the faith dimension at both these levels can play a role in participant well-being. Defining or categorizing programs by the presence of faith is the subject of ongoing debate (Ebaugh, Pipes, Chafetz, & Daniels, 2003; Smith & Sosin, 2001; Twombly, 2002).

Further, the particular interest in the context of the FBCI is whether the programs that have been the emphasis of the initiative thus far are effective. The FBCI was introduced in part to “level the playing field” in the competition for federal funding, acknowledging that historically many FBCOs had been unable to access these funds. Some FBCOs, such as Lutheran Family Services, Catholic Charities, and Jewish Family Services, found ways to successfully compete for federal funds. These large organizations had developed the internal capacity to deliver and manage services in a manner that satisfied federal funding restrictions. Thus, the focal interest now is on the effectiveness of programs that were not well represented in the previous federal funding picture.

This paper focuses on two key questions. First, to what extent are FBCO services effective? For the purposes here, *effectiveness* is defined as a program's ability to have participants show progress in achieving the desired outcomes (i.e., changes in knowledge, attitude, behavior, or status). For example, do children served by a mentoring program delivered by an FBCO show fewer behavioral problems and/or improved academic performance during or after the mentoring experience? The types of studies that are relevant to address this question are those that collect data on participants using some sort of outcomes measurement approach, which is often limited to pre-post assessments on a single group of participants (Bartowski, Call, Heaton, & Forste, 2007; Hangley & McClanahan, 2002; MacMaster et al., 2007; Rock, 2002; Roman, Wolff, Correa, & Buck, 2007). Although useful, these studies routinely lack outcome data from a comparison group that would provide a benchmark of what would have happened in the absence of the program. As such, these studies can only indicate whether participants changed during the time frame of the intervention and cannot address whether the change was a result of the program's efforts.

The second question addresses to what extent are FBCO services more effective than other approaches? For example, do clients served by different types of welfare-to-work programs (e.g., faith-based versus secular) show different levels of success in achieving the desired employment outcomes? The types of evidence that will inform this question are based on more rigorous evaluation designs in which meaningful comparisons can be made to a group of comparable individuals who received alternate services or usual care. This requirement is much more restrictive, and the number of studies that will achieve this standard is necessarily fewer than those that will address the first question. The most conclusive designs involve random assignment to groups or other quasi-experimental approaches to creating equivalent groups.

## **TAKING STOCK OF WHAT IS KNOWN ABOUT FAITH-BASED AND COMMUNITY-BASED SERVICES**

The present review comprises three steps: (1) a description of the previous reviews that have examined research in the FBCO domain, (2) a review of several study exemplars that illustrate particular strengths present in the literature, and (3) a review and synthesis of the existing comparative studies on FBCO effectiveness.

### **Learning from Previous Reviews**

Research reviews are an important building block in developing a research literature, in that they offer a periodic reflection on the state of the evidence base. In the FBCO arena, four previous reviews were identified and each is very briefly summarized in Table B-1 (see Appendix B).

Each of these focused on reviewing the characteristics and findings of research on FBCO services. These prior efforts sought to review the existing research base, describe its status, and summarize its contents narratively. None of the efforts attempted to quantitatively summarize evidence on program effects.

A critical and influential first review, entitled “Objective Hope,” was done by Johnson, Tompkins, and Webb in 2002. The report presents a review of nearly 800 studies, including a core group of 25 studies examining the effectiveness of FBCO services. The remainder of these studies related either to the relationship between religion and health outcomes or to the relationship between religion and other forms of well-being. Based on the quantity and quality of literature existing at the time on FBCO effectiveness, the authors concluded that although the overall body of work showed promising effects, most areas of FBCO service “have not been the subject of serious evaluation research” (p.21).

In 2003, the Roundtable on Religion and Social Policy released a review that focused on literature related to the scope and scale of faith-based and community-based activities (Scott, Montiel, Keyes-Williams, & Han, 2003). Although the primary intent was to summarize efforts at distinguishing between different types of FBCOs, describe their activities, and explore the amount of resources FBCOs were contributing through their work, the review included studies that also examined FBCO effectiveness. The authors grouped studies according to whether they focused on a specific locale or region, a multistate area, or a national perspective. The review included studies of FBCOs as well as congregations. The authors concluded that additional work was needed to document the specific coverage of FBCOs and to develop more complete data on the scope and scale of FBCO activities and comparative data from other service providers in a given geography. This review was focused on detecting the relative range of FBCO programming in the social service landscape, rather than on investigating the effectiveness of those programs.

In 2004, an additional review was published that specifically focused on studies of health programs in faith-based organizations and their effectiveness (DeHaven, Hunter, Wilder, Walton, & Berry, 2004). The types of health programs included in the review span the topics of general health, cardiovascular health, cancer, mental health, and nutrition. Approximately 43% of the programs involved the provision of health education or services in a church setting and an additional 25% were programs emerging from a congregation’s health ministry; the remaining 32% combined faith-based and secular elements. The authors reported that outcomes were measured in 28 (53%) of the studies and for each they reported whether the results were statistically significant. The authors concluded that faith-based health programs “can produce

positive effects” and reported that among the 16 studies that reported a statistical test, 15 showed a statistically significant difference favoring the intervention group compared with unserved populations (on outcomes such as decreased cholesterol and blood pressure and increased health and fruit/vegetable consumption). The authors also recommended strategies for increasing the number of effectiveness studies and their usefulness in health programming with diverse communities.

The most recent review, conducted by Ferguson, Wu, Spruijt-Metz, & Dyrness (2007), used the systematic review approach to examine how effectiveness has been defined in research on FBCOs. Using the keywords “faith-based” and “program effectiveness,” the authors searched a range of electronic databases as well as the Web sites of known institutional sponsors and producers of such research. The search ultimately identified 29 studies: 21 studies had a quantitative component and 8 studies used a comparative design. The authors discuss how effectiveness has been defined across these studies and offer a narrative summary of the findings within each of six outcome areas (e.g., health, criminal recidivism). They ultimately concluded, based on the limited number and quality of prior studies, that “the quality of findings from some previous evaluation studies on the effectiveness of faith-based programs remains questionable” (p. 272). They offered a number of useful recommendations for the field, such as working to broaden outcomes beyond the client level and more clearly describing the role of faith in program models.

Collectively, these four reviews reflect the state of the literature at the time of their completion as well as the needs for additional research. Combined, the reviews suggest three general observations. First, engagement in religious behaviors is convincingly associated with numerous indicators of positive health and well-being. The majority of these studies are correlational in nature; thus, they do not control for other factors. Second, there is a growing body of evidence that participants in FBCO programming do show improvement on identified outcomes over the course of their involvement with these programs. The limited number of comparative studies shows that FBCO-served populations may fare better in relation to comparison groups in some aspects. Third, the prior two observations do not provide sufficient evidence for documenting the benefit of FBCO programming. Rather, comparative studies using well-constructed reference groups are needed to more fully illuminate the effectiveness dimension.

### **Examining Study Exemplars**

Despite the limitations of the literature in this area, there are a number of noteworthy studies that have contributed considerably to the field of FBCO effectiveness research. The criteria for

identifying these exemplars were subjective; however, each study was selected to illustrate a particular contribution to the current knowledge base on FBCOs. The distinct contribution of each study will be described here. Three of these studies have been cited extensively in the literature for their findings and for the research methods they used. These three studies were also included in the review conducted by Ferguson et al. (2007). The fourth study highlighted was completed more recently in 2007 and is included because it used random assignment to groups.

The first study of note is Monsma & Soper's (2003, 2006) comparative evaluation of a range of welfare-to-work programs operated in Los Angeles. This study focused on comparing how welfare recipients served by programs operated by various types of providers differed in terms of their characteristics and their labor force outcomes. The authors selected 17 employment programs and then collected participant-level data from the selected programs. Although based on an observational design subject to selection bias concerns, this study offers a systematic and comprehensive look at a set of service providers all working toward the same set of outcomes, including two categories of faith-based employment programs.

Based on this analysis, the authors concluded that participants in the for-profit programs generally had fewer barriers to employment (e.g., criminal history, substance abuse, mental health issue), whereas participants in the programs with the more intensive faith elements generally had more barriers. The study found that the faith-based programs were the least successful in placing unemployed participants into jobs by 6 and 12 months, but were somewhat more effective in helping employed participants retain their employment once placed.

A second study is the evaluation of Amachi, a model mentoring program for children of prisoners in Philadelphia. The Amachi program is based on a partnership between secular and faith-based organizations (Jucovy, 2003). This study focused on examining how the model was implemented and how the mentoring relationships fared over time. Amachi adopted the outcome model used by Big Brothers–Big Sisters (BBBS), which previously found that positive results for mentees begin to occur after 12 months of engagement. Using this benchmark, 62% of Amachi matches were active 12 months or longer and exceeded engagement in the generic BBBS program (46%). This study demonstrated how faith-based services can adopt secular outcomes measurement approaches from the same program domain.

A third notable study involves the InnerChange Freedom Initiative (IFI), a faith-based prerelease program for prisoners operated in Texas (Johnson & Larson, 2003). The intent of this study was to assess the implementation of the program and measure participant outcomes. The program involves in-prison, Bible-based programming and 6 to 12 months of aftercare once the

participant is paroled. The authors implemented a matched group design, whereby outcomes for IFI participants (and graduates) could be contrasted with three comparison groups of prisoners who were similar to the IFI group in important ways.

Overall, the study found that IFI participants showed equivalent or slightly more recidivism than the three comparison groups on measures of rearrest and reincarceration within 2 years of release. However, the authors reported that program *graduates* demonstrated dramatically lower recidivism rates over 2 years compared with noncompleters—17% versus 50% on rearrest and 8% versus 36% on reincarceration, respectively (Johnson & Larson, 2003). Although the experience of graduates does not generalize to all program participants, the contrast with noncompleters may be important from a programmatic stance. Examining the characteristics of those who complete may suggest whom the program engages most successfully. For example, in the case of IFI, program graduation was more frequent among Hispanic men, men over 35 years of age, and men judged to be at low risk, compared with the other categories of participants.

A fourth exemplary study involved a faith-based intervention with cocaine-addicted women (Stahler, Kirby, & Kerwin, 2007). Although based on a very small sample size of 18, the study showed how random assignment could be used effectively to compare treatment alternatives. Following a recruitment and screening procedure, the 18 study volunteers were randomly assigned to the two groups. Both groups of women received the supervised residential treatment component of the program. In addition, the intervention group received faith-based workshops and mentoring (called the Bridges program), while the control group was exposed to other group programming of a secular nature. The study showed significantly better program retention among the Bridges group, as well as higher rates of drug abstinence at 6 months postintake based on urine samples (75% versus 30%). However, the differences between the groups were not statistically significant at 3 months postintake.

### ***Summarizing the Available Studies***

A final stage of this work is to engage in a synthesis of the existing comparative research. The task of reviewing the available evidence on the effectiveness of FBCOs is a process of both inclusion and exclusion. To clarify what is known on this topic, decisions were made as to what evidence is relevant and what studies can inform the process. Three categories of research were not deemed sufficiently relevant to the present policy discussion and were not reviewed as part of this effort. The three domains—studies of religiosity, studies of programs targeting church congregations, and studies that are descriptive in nature or only use qualitative methods—are



described in Appendix C. These study domains are important, but they are less relevant to answering more narrow questions about program effectiveness.

## METHODS

The study methods used here are based on the accepted standards for research synthesis approaches (Lipsey & Wilson, 2001). The framework of the research synthesis procedure can be cast in terms of six interrelated steps: (1) specification of the questions of interest, (2) specification of the research domain, (3) retrieval of empirical studies, (4) review and coding of studies, (5) statistical aggregation of research results, and (6) presentation and interpretation of findings (Cordray & Fischer, 1994).

With the primary focus of the review effort delineated, a search was conducted to locate as many candidate studies as possible for inclusion in the review. The search had two main components. First, the existing research reviews were examined and the studies included in each of these were culled for candidates for the present review (DeHaven et al., 2004; Ferguson et al., 2007; Johnson et al., 2002; Scott et al., 2003). The availability of prior reviews is a major asset and these offer a key starting point for the search procedure.

Second, a search was conducted of 10 large electronic databases containing publications and reports in the social and behavioral sciences.<sup>1</sup> The disciplines represented in these databases include anthropology, economics, law, medicine, nursing, policy studies, psychology, social work, and sociology. The search was focused using the search terms “faith-based,” “community,” and “evaluation” together, as well as “faith-based” and “outcome” together. Studies targeted for inclusion had all of the following attributes: (1) reported on the evaluation of a program delivered by a faith-based or community-based organization,<sup>2</sup> (2) reported quantitative data on participant outcomes relevant to the intervention (e.g., employment, substance abuse, criminal activity), (3) used a comparative study design involving a group of program participants and a group of individuals who participated in an alternate program or no special programming, (4) involved the evaluation of a human service programming effort similar in substantive nature to those targeted by the FBCI, (5) involved programming that was directed to a general population of individuals rather than a group composed of members of a religious congregation, and (6) was reported in 1990 or later. Of particular interest were studies completed after 2004 because the existing reviews by Johnson et al. (2002) and Ferguson et al. (2007) had identified the vast majority of potential studies completed through 2004. The focus of this synthesis is on studies that examine the outcomes of faith-based programs, often in comparison to secular programs with the same objectives.

The present review differs from the prior review efforts in two important ways. First, this review specifically targets only those studies that made use of a comparative design in assessing quantitative outcomes. The goal here was to restrict attention to the potentially most rigorous studies that could address the issue of effectiveness. For example, a group of studies included in the review by Ferguson et al. (2007) were excluded here, including eight studies that used only qualitative methods and an additional four studies that relied on a single-group design. An additional 10 single-group studies in the review by Johnson et al. (2002) were also excluded. Second, the present review extends the search window by three additional years to 2007. This additional period is quite important, especially in a field that has seen such dramatic growth and attention over the past few years.

Based on the search procedures, a total of 92 independent citations were generated. Once identified, copies of publications and reports were secured and examined for eligibility by the author. In this process, particular emphasis was given to the type of research design used in the study, distinguishing between studies that used a single-group pre-post or related design and those that were truly comparative in nature (i.e., using two or more groups). A supplemental search technique was also applied by reviewing the citations and footnotes of all retrieved studies. Studies that were ultimately deemed to meet the inclusion criteria were then formally reviewed.

The review of the existing literature identified a meaningful core of studies that examined the outcomes of FBCO programming. However, only 18 of the 92 quantitative outcomes studies identified used a comparative research design. These studies are summarized in Table D-1 (see Appendix D). Among the 18 studies identified, 13 had been included in at least one of the prior reviews. For each study, the table highlights the target population, the study design and sample size, a brief statement of the intervention, the outcomes measures, the basic findings, whether subgroup analyses were conducted, and any relevant effect size information. The 18 studies span six distinct target populations, with multiple studies focused on prisoners and former prisoners (8), welfare clients (4), substance abusers (2), the elderly (2), and additional studies of children of prisoners and Latino women. Sample sizes vary dramatically, as do the procedures used to construct a comparison group for the studies. The key outcomes of interest are specific to the substantive focus of the programs. Although there is some consistency within program domains (e.g., recidivism among prisoner programs), the time frames of the outcomes vary. Finally, subgroup analyses were examined in over two thirds of the studies, but these varied in scope and focus.

For each of the studies, the reported outcomes data were collected and used to compute standardized mean effects on the key program outcomes identified by the study's authors. The effect size provides a standardized metric for comparing across studies and outcomes. Of the 18 studies, 14 compared faith-based programming of some type with either a specific secular program or generally available services delivered to the same target population. The remaining four studies compared a faith-based program with the experience of individuals served by no specific program.

The unweighted mean effect size of all 18 studies is .216 and ranges from a high of .736 to a low of  $-.807$ . Effect sizes with a positive value indicate that the intervention group had greater success on the outcomes, while negative values show better success in the comparison group. The unweighted effect size does not take into account the differing sample sizes underlying the studies, as studies based on larger samples tend to provide more precise estimates and should be given more weight in this approach. When weighted for the study sample size (and after removing one study deemed to be a statistical outlier), the mean effect size is .205. Figure E-1 (see Appendix E) presents a summary of distribution of effect sizes. The overall mean effect size would be judged in the small range according to the standard offered by Cohen (1988) and in the bottom 25% of effects according to Lipsey and Wilson (2001). Another interpretation of an effect size of .2 is that it equates to a 10% difference on a success measure between the treatment and comparison groups (e.g., 55% versus 45%).

Two of the larger subgroups of studies were also examined. The eight studies of interventions with prisoners and former prisoners yielded a weighted mean effect size of .138, again in the small range. The three studies of welfare-to-work programming (after removing the one statistical outlier) yielded a weighted mean effect size of .503. This effect is in the moderate range and is equivalent to a 24% difference in the success rates between groups (e.g., 62% versus 38%).

The set of available studies that use a comparative design and report quantitative outcomes in sufficient detail is relatively small and is only able to support a preliminary quantitative synthesis of findings. As such, at present, the findings presented should be seen as illustrative of the kinds of beneficial impacts that FBCOs can produce across a range of substantive domains working with differing target populations. On average, the presence of faith in these programs, along with the other characteristics of the interventions, appears to result in a modest effect on the outcomes of interest over those observed among comparison populations.

## **FINDINGS AND IMPLICATIONS**

Since its launch in 2001, the FBCI has dramatically increased the role of smaller FBCOs in the delivery of federally funded social services (White House, 2008). Concurrently, there has been expanded interest in the ability of FBCOs to document their outcomes, and in identifying high-quality research to test the effectiveness of services delivered by FBCOs. Despite considerable effort and progress, the existing evidence base remains limited. A preliminary quantitative synthesis shows that the overall effect of FBCO programs, although modest in size, demonstrates that these programs tend to produce somewhat better outcomes compared with usual services, secular services, or no special programming. More data are needed to confirm this finding, as are data to investigate the effectiveness of specific categories of programs for defined target populations. In order to further contribute to existing knowledge in this domain, efforts should be continued in at least three areas, as described below.

### **Outcomes Measurement and FBCO Capacity**

Several challenges to evaluating the effectiveness of FBCO services stem from the fact that FBCOs targeted since 2001 tend to be small nonprofits with limited capacity for ongoing data collection or systematic research studies (Fischer, 2004). The growth of outcomes measurement has spurred a major shift in the way nonprofits view their work and the way they communicate their work to their funders, clients, and other stakeholders (Fischer, 2001; Hatry, Van Houten, Plantz, & Greenway, 1996). The Compassion Capital Fund National Resource Center (2005) has recognized the value of this approach and has produced a manual on outcomes measurement for use by intermediary organizations assisting FBCOs to build capacity.

The available evidence suggests that because of their limited size and relative inexperience with outcomes measurement, many FBCOs need specific assistance to develop capacity to collect, manage, and analyze their data. Johnson et al. (2002) reported that FBCO supporters often cite exceptionally high rates of success for programs, but that “closer examination of these accounts...tends to reveal mere simple summary statistics based on in-house data compiled by the religious organizations and ministries themselves” (p. 15). FBCOs should strive to collect more complete and accurate data and present it in a methodologically rigorous and neutral fashion. Some researchers have called for federal policy makers to emphasize the need for more and better evaluation of FBCO services, going so far as to suggest a requirement for clear logic models along with financial support for evaluation efforts (Mears, Roman, Wolff, & Buck, 2006).

A central implication of the capacity issue is the imperative of addressing the developmental needs of FBCOs. The recognition that FBCOs require specialized assistance in fully developing and assessing their programs has resulted in the funding of intermediary organizations to help build the capacity of FBCOs (Sherman, 2002). For example, of the Compassionate Capital Fund monies initially appropriated, \$25 million (83%) was for intermediary organizations to aid FBCOs “to replicate or expand best practices and model programs in targeted areas” (Sherman, 2002). As the capacity of FBCOs is better understood, there can be better planning to address their capacity-related needs (Clerkin & Gronbjerg, 2007; Leake et al., 2007).

There is a strong emphasis on working to increase the capacity of FBCOs through promoting internal development and external support via intermediary organizations (Fink & Branch, 2005; Sherman, 2006). In fact, the strategy is now recognized as one of the key innovations of the FBCI (White House, 2008). Because organizational capacity is inextricably linked to an organization’s ability to document its outcomes and take part in more rigorous research, investments in FBCO capacity will facilitate further development of the research literature as well.

Outcomes measurement may also prove to be an empowerment mechanism for FBCOs. Fagan, Horn, Edwards, Woods, and Caprara (2007) suggest that “outcome-based evaluation has the potential to engender a revolution of increased effectiveness in the faith community and debunk skeptics’ claim that faith-based programs are only about ‘feel good’ results rather than producing solid and measurable impacts” (p. 1).

### **Rigorous Evaluation**

The ongoing dialogue over the FBCI involves a range of concerns but continues to include a heavy emphasis on the effectiveness of FBCO services compared with both their secular counterparts and/or conventional services (DiIulio, 2002). The most promising avenues for responding to the data needs regarding FBCOs are through improving and expanding data collection practices and fielding more rigorous comparative studies to address issues of effectiveness. In the evaluation of any human service intervention, it is widely accepted that the use of randomized, well-implemented research designs leads to the most credible assessments of program impact. Experimental and quasi-experimental research designs have the distinct advantage of eliminating the role of a range of plausible intervening factors that could compete with the program in explaining impacts (Cook & Campbell, 1979). To date, only one study of faith-based services using a randomized design has emerged (Stahler et al., 2007).

Federally funded research on FBCI-related programming may hold the most promise in regard to expanding the use of more rigorous designs. As highlighted in a recent White House report in a chapter called “Measurement Matters,” no less than 20 studies of FBCI-related programs are currently underway by nine federal agencies (White House, 2008). These should be monitored to gauge what they may be revealing about effective programming strategies. The evaluation research community and funders of such efforts should expand the discussion of how to bring more resources, both technical and financial, to bear on the evaluation needs of the FBCO sector. These efforts will necessarily involve governmental and contractual oversight by funders to monitor the execution of FBCI activities, but must also get at the intended participant outcomes (Government Accountability Office, 2006). In addition, although more rigorous research is needed, this should not preclude further important advancements based on qualitative and observational research approaches. These research domains help bring clarity to the context of these programs and the lives of participants, and offer unique and richer understanding of program delivery and effectiveness.

### **Operationalizing Faith**

An area of great interest and debate has been in characterizing the nature of faith-based programs (e.g., Monsma & Mounts, 2002; Sider & Unruh, 2004; Smith & Sosin, 2001; Working Group on Human Needs, 2002). For example, the Working Group (2003) defines an FBCO as “any entity that is self-identified as motivated by or founded on religious conviction” (p.2). The ability to assess the relative degree of faith intensity of a social service program is central to clarifying the program’s theory, logic, and ultimately the key outcomes. If the role of faith is a key ingredient in the expected success of the faith-based programs, then it is essential to better understand and measure its presence (Fischer, 2004). Faith can be both a matter of the context or environment of programs as well as part of the intervention itself, and as yet there are very limited data on this distinction.

## **CONCLUSION**

Despite the substantial growth in the funding available to faith-based and community-based organizations over the past decade, the field of research on FBCOs remains very young and underdeveloped. As recently as 2002, the U.S. General Accounting Office concluded that the literature “provides no information on which to assess the effectiveness of FBCOs as providers of social service” (p. 17). Although advances have been made and a productive dialogue is underway, the extent of the existing evidence base is insufficient as a guide for program planning and enhancement. Overall, because of the relative youth of the FBCO research field, there is a

lack of systematic data on FBCO services and their effectiveness. As indicated earlier, much of the existing research on FBCO services is descriptive in nature, with a focus on programmatic models, delivery styles, and funding streams of FBCO services (Independent Sector, 2003). However, the research in the field has expanded over time (1998 to 2007) and now includes a number of key empirical studies as well as important efforts to distill what is known about the effectiveness of existing programs. The field needs to move forward with an agenda of establishing data systems for the purposes of accountability, program improvement, and demonstrating effectiveness.

### **ACKNOWLEDGMENTS**

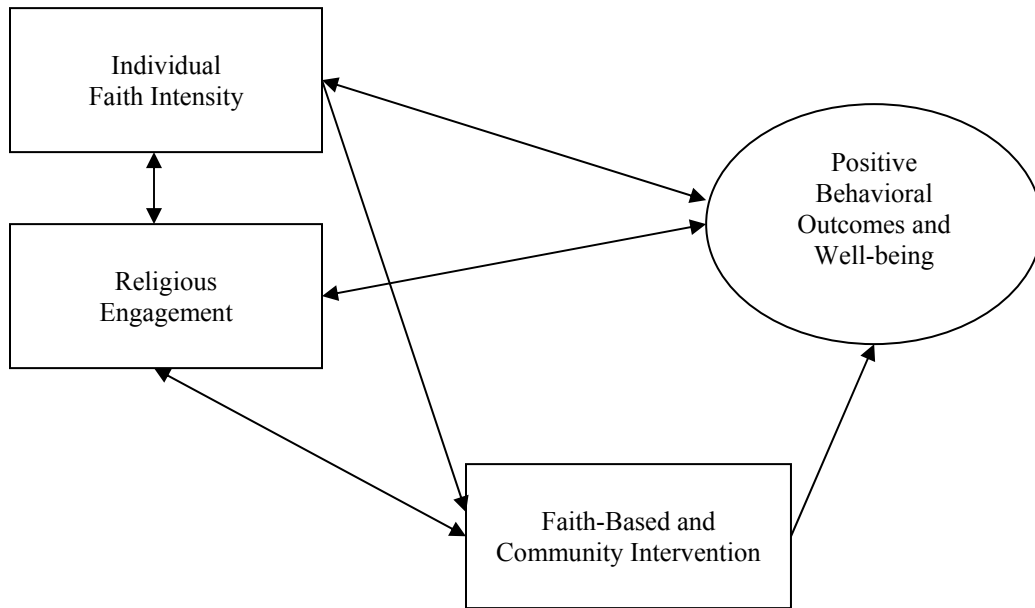
The author wishes to thank Grace Dyrness, Kristin Ferguson, and Mark Joseph for helpful comments provided on an earlier draft, as well as Diwakar Vadapalli and Monica Craven for assistance in preparing the manuscript.

### **NOTES**

1. The databases included PsychINFO, Social Science Index, Social Work Abstracts, BiblioLine, Lexis-Nexis, MEDLINE, Dissertation and Theses Abstracts, Nursing and Allied Health Sources, Health Source, and Arts and Humanities.
2. The focus here is on what Chaves (2002) calls “religious nonprofits,” defined as “religious organizations working in nonreligious functional fields includ[ing] the wide variety of religious organizations doing virtually everything secular nonprofit organizations do” (p.1524).

## APPENDIX A

**Figure A-1. Linkages Between Faith and Wellness**





## APPENDIX B

**Table B-1. Reviews of Research Related to the Faith-Based and Community Initiative**

| Authors & Title   | Year | Number of all studies reviewed | Number of quantitative studies included | Date range of studies included | Focus of studies included   |
|---|------|--------------------------------|---|--------------------------------|---|
| Johnson, Tompkins, & Webb— <i>Objective Hope</i>  | 2002 | 766                            | 25                                      | 1933–2002                      | Relationship between religion and health and well-being outcomes; efficacy of FBCOs |
| Scott, Montiel, Keyes-Williams, & Han— <i>The Scope and Scale of Faith-Based Social Services</i>                                      | 2003 | 75                             | Not a specific focus                    | 1992–2003                      | Documenting the scope and scale of FBO activities documented in studies             |
| DeHaven, Hunter, Wilder, Walton, & Berry— <i>Health Programs in Faith-Based Organizations: Are They Effective?</i>                    | 2004 | 53                             | 28                                      | 1990–2000                      | Health programs delivered in churches or by health ministries                       |
| Ferguson, Wu, Spruijt-Metz, & Dyrness— <i>Outcomes Evaluation in Faith-Based Social Services: Are We Evaluating Faith Accurately?</i> | 2007 | 29                             | 29                                      | 1987–2004                      | Program effectiveness and faith-based organizations                                 |

Note. FBCO=Faith-Based and Community Organizations; FBO=Faith-Based Organization

## APPENDIX C

### Systematic Review Methods

#### *Study Domains Not Included in Present Review*

1. *Studies of the relationship between religiosity and the status of individuals.* These studies examine the relationship between religious behaviors (e.g., church attendance), beliefs, and attitudes, and the occurrence of a range of positive and negative behaviors. Research on what Johnson et al. (2002) termed “organic” religion, demonstrates that religious participation and belief are related to a wide range of positive social and health outcomes (Johnson, 2002; Johnson & Siegel, 2003; Powell, Shahabi, & Thoresen, 2003; Wilcox, 2002). There is also a

strong correlational link showing that youths with more religious involvement show better academic progress (Regnerus, 2002), less juvenile delinquency (Johnson, De Li, Larson, & McCullough, 2000), less criminal activity (Baier & Wright, 2001), and are more likely to engage in a range of healthy behaviors (Wallace, 2002). These studies provide important context, but they do not speak to the effects of particular program initiatives.

2. *Studies of programs targeting church congregations.* These so-called “faith-placed” efforts use the church setting as a venue for delivering an intervention, targeting members of a congregation for programs often delivered by other members. These studies often involve services delivered to the members of a congregation, not to a general target population in need. The challenge is that congregation members are often considered the active agent of service provision rather than the recipients of service. Congregations are excluded because they do not reflect the primary theoretical frame underlying the FBCI that involves offering services to a community population that is not simply the membership of a faith community. For example, studies have examined the effectiveness of offering church-based breast and cervical cancer screening interventions (Matthews, Berrios, Darnell, & Calhoun, 2006), automobile restraint use (Falcone, Brentley, Ricketts, Allen & Garcia, 2006), breast self-examination and mammography promotion (Erwin, Spatz, Stotts, & Hollenberg, 1999), fruit and vegetable consumption (Campbell et al., 1999), and smoking cessation (Schorling et al., 1997). Although it is clear that the vast majority of congregations do self-identify as offering one or more types of social service programming (e.g., Cnaan & Boddie, 2001), these programs are often quite distinct from services offered by FBCOs. Some scholars (e.g., Jeavons, 2004) have argued that, and for this and other conceptual reasons these studies are not reviewed.
3. *Studies that are descriptive in nature or use only qualitative methods.* Numerous studies have examined the approach and theory underlying FBCO interventions or have focused exclusively on assessing the implementation of these services. For example, Hodge and Pittman (2003) examined the characteristics of a sample of faith-based drug and alcohol treatment providers in Texas. This type of work is critically important for a number of reasons, including the essential clarification of program logic and program delivery models. This clarity will aid in the effective evaluation of these programs and serve to inform program development efforts across the board (Fischer & Stelter, 2006). However, such studies do not measure program benefit or document participant improvement. While it is true that FBCO social service programs contain many of the same primary programmatic

elements as secular programs, these faith-based elements should be clearly understood, quantified, and assessed as the field moves forward.

### ***Effect Size Calculations***

The formulation of the effect size metric is based on Cohen's (1969) original notion of standardized mean difference (d statistic) between two groups on an outcome of interest. In a treatment-effectiveness synthesis, the effect size is characteristically calculated as the raw difference in means for two groups divided by a measure of variation (e.g., the standard deviation) of the outcome distribution. Among these studies, group-level standard deviations were rarely reported, so the pooled variances were estimated using a procedure based on sample size (Raudenbush, 1994). The effect size can be calculated indirectly from reported significance test information for a given outcome. Other related transformations and manipulations are available for converting various reported significance and outcome information into the d statistic (Rosenthal, 1994).

Two calculations were necessary for the core meta-analytic procedure. First, a measure of impact (e.g., difference between treatment and control groups) was needed as the basic unit of the synthesis (d). Second, a sampling error variance estimate (v) was needed for each effect size estimate. The majority of the studies provided success rates for both groups, and these data were converted to an effect size using Cohen's h statistic, an arcsine transformation (Lipsey & Wilson, 2001). Estimated effect sizes were computed for the remaining studies using whatever statistical data the authors presented in the original report. In the case of studies with multiple outcomes, the average effect size of the reported measures is presented in the Table D-1 (see Appendix D).

### ***Limitations***

The quantitative synthesis procedure (i.e., meta-analysis) is a useful tool but it does have a number of limitations. The external validity of research synthesis is important for the purpose of generalizing the review findings beyond the current set of studies. All reviews are threatened by bias due to the selectivity of the studies that are included. One particular concern is that meta-analysis is "dependent on the findings that researchers report" and bias will result if there are systematic differences among the results of research depending on whether they are published or not. This concern arises from the idea that many study results remain in "file-drawers" (Rosenthal, 1979) because their authors did not seek publication or were rejected, presumably because the results showed nonsignificant effects. As a result, the research retrieved by a meta-analysis may have a systematic bias. The present review handled the "file-drawer" threat by making an extensive search of the published and unpublished outlets for research on FBCOs.

Ultimately, only 3 of the 18 studies were retrieved from peer-reviewed sources. As such, the concern about a bias due to the file-drawer effect is substantially reduced.

An additional limitation is that quantitative reviews are dependent on the focus, quantity, and quality of the component studies available. As such, a review can only examine areas of research that have been the focus of attention in a sufficient number of individual studies. Also, the limitations of the individual studies in regard to data quality, reliability of measures, sample attrition, etc., are also a factor in the results of the review. As such, a review is ultimately only as good as the studies it comprises. The present review attempts to focus on the relatively higher quality studies available by restricting inclusion to those studies that use a comparative research design of some type (versus single-group designs).

**APPENDIX D**

**Table D-1. Comparative Outcomes Studies of FBCO Services**

| <b>Study</b>   | <b>Target Population</b>                               | <b>Design and Sample Size</b>                               | <b>Intervention Description</b>  | <b>Outcomes Measures</b>                      | <b>Findings</b>   | <b>Subgroup Analyses</b>                              | <b>Mean Effect Size/95% CI</b> |
|--|--|---|--|---|---|---|--------------------------------|
| Bicknese (1999) – Teen Challenge                                 | Substance users (TC graduates in PA, MO, and CA)       | Tx (59) vs. C (118); matched controls from national sample  | Tx – Bible-based classes plus employment<br>C – short-term inpatient program             | Substance abuse, employment, criminality      | FBO group significantly more abstinent from drugs/alcohol (71.2% vs. 55.1%), more held full-time jobs (89.8% vs. 41.4%), fewer with arrest (7.0% vs. 17.0%) | By type of substance used                             | .581<br>(.558 – .604)          |
| Campbell (2004) – Charitable Choice                              | Welfare clients in North Carolina                      | Tx (1,320) vs. C (10,862); Workforce Investment Act clients | Tx – faith-based employment services<br>C – governmental employment services             | Employment at program exit                    | Faith-based programs showed less employment success compared with governmental centers (40.9% vs. 79.2%)  | By type of provider                                   | –.807<br>(–.811 – –.809)       |
| Deb & Jones (2003) – FaithWorks                                  | Welfare clients, statewide in Indiana                  | Tx (2,930) vs. C (2,397); statistical controls              | Tx – FaithWorks job training<br>C – secular job training                                 | Job placement, wages, hours, health insurance | Similar placement and wage rates; FBO clients work significantly fewer hours and fewer acquire jobs with health insurance                                   | By gender, race, and high-school diploma status       | .550<br>(.549 – .551)          |
| Farley & Hackman (2006), Farley & McClanahan (2007) - Ready4Work | Former prisoners in 9 sites (6 faith-based, 3 secular) | Tx (2,374) vs. C (national sample – BJS)                    | Tx – employment readiness, job placement, mentoring<br>C – usual services in communities | Recidivism at 6 months and 12 months          | Participants had lower recidivism at 6 months (1.9% vs. 5.0%) and at 12 months (5.0% vs. 10.4%)   | 18- to 34-year-old African-American nonviolent felons | .190<br>(.189 – .191)          |

(continued)

**Table D-1. Comparative Outcomes Studies of FBCO Services (continued)**

200

| Study   | Target Population                                      | Design and Sample Size   | Intervention Description   | Outcomes Measures                       | Findings  | Subgroup Analyses                             | Mean Effect Size/95% CI  |
|---|--|--|--|---|---|---|--------------------------|
| Florida Dept of Corrections (2000) – Kairos Horizons                | Prisoners in Florida                                   | Tx (67) vs. C (741); in Tomoka Prison                          | Tx – prisoners in faith-based dorm<br>C – prisoners in other dorms                       | Disciplinary reports                    | Prisoners in faith-based program had fewer disciplinary reports (8.9% vs. 16.7%)  | Completers vs noncompleters                   | .236<br>(.227 – .245)    |
| Fox, Stein, Gonzalez, Farrenkopf, & Dellinger (1998)                | Latino women in two communities in southern California | Tx (176) vs. C (126); samples in two communities 1990 and 1998 | Tx – faith-based educational programming and mammography services<br>C – usual services  | Mammography use in previous year        | Differences in mammography awareness (+19% vs. +12%); significant improvement in Tx community in receipt of mammography (27% vs. 24%) | Spanish speaking vs non-Spanish speaking      | .174<br>(.161 – .187)    |
| Johnson (2002) – Humaita  | Prisoners in two prisons in Brazil                     | Tx (247) vs. C (148); matched sites                            | Tx – Prison staffed by church volunteers<br>C – Prison operated by nonprofit             | Recidivism over 3 years                 | FBO prisoners had significantly lower recidivism (16% vs. 36%)  | High-risk vs low-risk                         | .464<br>(.454 – .474)    |
| Johnson & Larson (2003) – InnerChange Freedom Initiative            | Prisoners in Texas                                     | Tx (177) vs. C (1,754); matched group in Texas                 | Tx -In-prison bible-based programming and 6-12 months of aftercare<br>C – usual services | Rearrest and incarceration over 2 years | No significant difference on rearrest (36.2% vs. 35%) or incarceration (24.3% vs. 20.3%)  | Completers vs noncompleters                   | -.061<br>(-.065 – -.057) |
| Johnson, Larson, & Pitts (1997); Johnson (2004) – Prison Fellowship | Prisoners, in four New York prisons                    | Tx (201) vs. C (201); matched group                            | Tx – Prison Fellowship<br>C – no intervention  | Recidivism at 2 years, 3 years, 8 years | No significant differences between two main groups at 2 years (44% vs. 43%), 3 years (53% vs. 57%), or 8 years (68% vs. 73%)          | Level of program participation, level of risk | .057<br>(.047 – .067)    |

(continued)

**Table D-1. Comparative Outcomes Studies of FBCO Services (continued)**

| Study  | Target Population                               | Design and Sample Size  | Intervention Description  | Outcomes Measures                              | Findings   | Subgroup Analyses         | Mean Effect Size/95% CI  |
|--|---|---|---|--|--|---------------------------|--------------------------|
| Jucovy (2003) – Amachi   | Children of prisoners in Philadelphia           | Tx (399) vs. universe of BBBS mentor matches                                    | Tx – Volunteer mentors from churches<br>C – Standard mentors  | Mentor relationship lasting more than one year | Program had higher rate of matches lasting one year or more (62% vs. 46%)  | n/a                       | .322<br>(.317 – .327)    |
| LaVigne, Brazzell, & Small (2007) – Faith & Character-based Institutions | Prisoners in Florida                            | Tx (289) vs. C (289); matched group from waitlist and general prison population | Tx – variety of faith-based programs delivered by nonprofits<br>C – no special programming                          | Recidivism over 6 and 12 months                | Significantly lower recidivism for males at 6 months (0% vs. 2.1%) but not at 12 months (1.8% vs. 2.4%); nonsignificant difference for females at 6 months (0% vs. 1%) and 12 months (1.9% vs. 6.5%) | Males vs females          | .167<br>(.160 – .174)    |
| Modesto (2006) – Welfare-to-Work   | Welfare recipients in 3 North Carolina counties | Tx (102) vs. C (78); matched  | Tx – faith-based program plus community college program<br>C – secular program +community college                   | Earnings, poverty, employment                  | No significant differences between the groups on the measures available  | County of residence, race | .383<br>(.361 – .405)    |
| Monsma & Soper (2003, 2006); Monsma (2006) – Welfare-to-Work             | Welfare recipients in Los Angeles               | Tx (102) vs. C1 (141), C2 (113), C3 (80); clients served at 17 agencies         | Tx – welfare-to-work program from FBO<br>C – welfare-to-work program from government, for-profit, or secular agency | Employment (6/12 months), wages, TANF receipt  | FBO programs more successful in helping employed clients retain a job (~90% vs. 61%–77%) but less successful in getting unemployed clients employed (~24% vs. 42%–60%)                               | By 5 provider types       | -.015<br>(-.026 – -.004) |

(continued)

**Table D-1. Comparative Outcomes Studies of FBCO Services (continued)**

| Study  | Target Population                       | Design and Sample Size                              | Intervention Description   | Outcomes Measures                               | Findings  | Subgroup Analyses   | Mean Effect Size/95% CI |
|--|---|---|--|---|---|---|-------------------------|
| O'Connor, Su, Ryan, Parikh, & Alexander (1997) – Transition of Prisoners (TOP) | Former prisoners in Detroit             | Tx (95) vs. C1 (88) waitlist, C2 (85) matched group | Tx – church-based mentor and support<br>C – no intervention  | Recidivism over 3 years                         | FBO program had lower rate of return to prison for escape (33% vs. 43%, 57%); no difference in recidivism due to parole violation or new crime (33% vs. 21%, 34%) | By program participation level, education level, prior felony | .254<br>(.239 – .269)   |
| Ragan (2004a) – Nursing Homes  | Nursing home clients, national sample   | Tx (948) vs. C (15,342); nursing homes              | Tx – religiously affiliated homes<br>C – secular homes   | Inspection deficiencies, complaint deficiencies | FBO homes had significantly lower rates of inspection deficiencies (4.40 vs. 5.93) and complaint deficiencies (1.46 vs. 3.44)                                     | n/a   | .193<br>(.192 – .194)   |
| Ragan (2004b) – Home Health Agencies   | Home health clients, national sample    | Tx (445) vs. C (6,723); home health agencies        | Tx – religiously affiliated home health agencies<br>C – secular agencies   | 11 measures of patient outcomes                 | 8 of 11 measures showed that patients served by religiously affiliated agencies improved significantly more   | n/a   | .031<br>(.030 – .032)   |
| Stahler, Kirby, & Kerwin (2007) – Bridges program                              | Cocaine-dependent women in Philadelphia | Tx (8) vs. C (10); random assignment                | Tx – residential program, faith-based mentoring and group activities<br>C – residential program and secular group activities | Substance use at 3 and 6 months                 | No significant difference in abstinence at 3 months (88% vs. 66%), but at 6 months FBO program had significantly higher rates of abstinence (75% vs. 30%)         | n/a   | .736<br>(.517 – .955)   |

(continued)



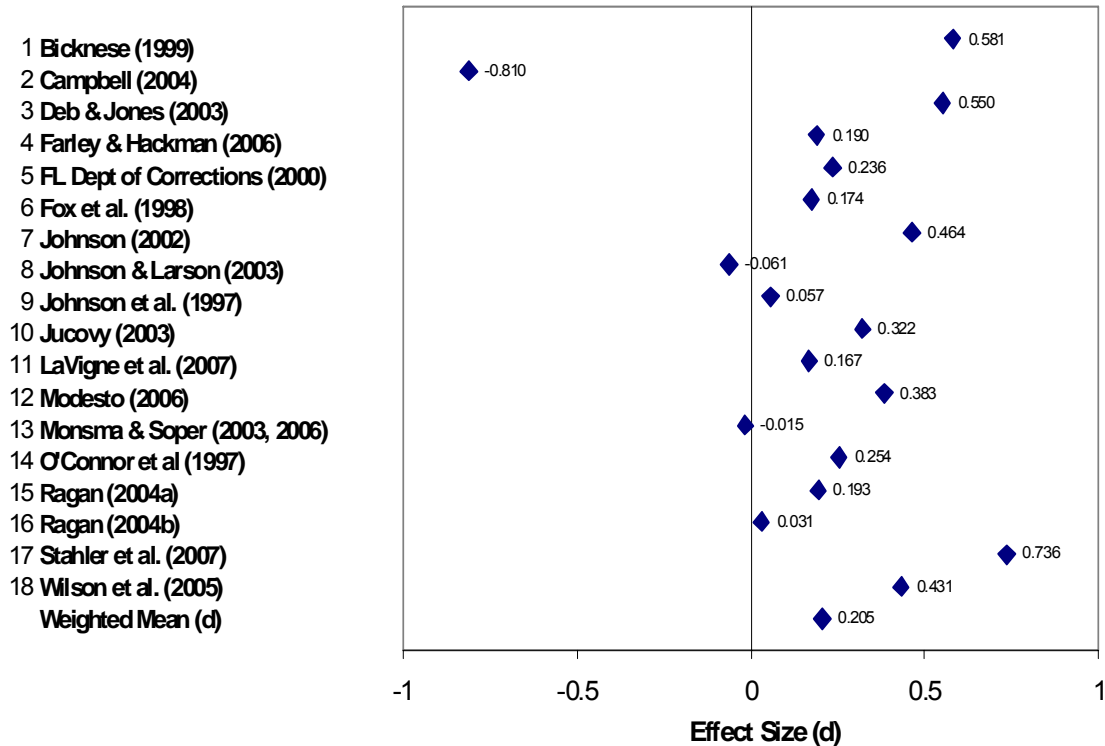
**Table D-1. Comparative Outcomes Studies of FBCO Services (continued)**

| <b>Study</b>   | <b>Target Population</b>         | <b>Design and Sample Size</b>     | <b>Intervention Description</b>  | <b>Outcomes Measures</b>                     | <b>Findings</b>  | <b>Subgroup Analyses</b> | <b>Mean Effect Size/95% CI</b> |
|--|----------------------------------|-----------------------------------|--|--|--|--------------------------|--------------------------------|
| Wilson, Picheca, & Prinzo (2005) – Circle of Support and Accountability (COSA) | Sex offenders in Ontario, Canada | Tx (60) vs. C (60); matched group | Tx – Participants assigned to 5 community volunteers who form a support group<br>C – No intervention | Recidivism – sexual offense, violent offense | Program group had significantly fewer sexual offenses (5.0% vs. 16.7%) and violent offenses (15.0% vs. 35.0%) over 4.5 years | n/a                      | .431<br>(.398 – .464)          |

Note: In the design and sample and intervention description columns, Tx refers to the treatment condition and C refers to the comparison condition. The Ragan (2004) report contains two distinct studies, which are listed separately here as (2004a) and (2004b).

## APPENDIX E

**Figure E-1. Effect Sizes for Studies Included in the Analysis**



Note: The Ragan (2004) report contains two distinct studies, which are listed separately here as (2004a) and (2004b).

## REFERENCES

- Baier, C. J., & Wright, B. R. E. (2001). If you love me, keep my commandments: A meta-analysis of the effect of religion on crime. *Journal of Research in Crime and Delinquency*, 38(1), 3–21.
- Bartowski, J. P., Call, V. R. A., Heaton, T. B., & Forste, R. (2007). Religion, job readiness, and employment outcomes: The case of Latter-Day Saint Employment Resource Services. *Research on Social Work Practice*, 17(2), 188–198.
- Berk, R. A. (1983). An introduction to sample selection bias in sociological data. *American Sociological Review*, 48, 386–398.
- Bicknese, A. T. (1999). The teen challenge drug treatment program in comparative perspective. Unpublished doctoral dissertation, Evanston, IL: Northwestern University.

- Branch, A.Y. (2002). *Faith and action: Implementation of the national faith-based initiative for high risk youth*. Philadelphia: Public/Private Ventures. Retrieved December 10, 2004, from <http://www.ppv.org>
- Camp, S. D., Klein-Saffran, J., Kwon, O., Daggett, D. M., & Joseph, V. (2006). An exploration into participation in a faith-based prison program. *Criminology & Public Policy*, 5(3), 529–550.
- Campbell, D. (2004). *CFBO roles in serving the hard-to-employ: A government-led partnership to realize public value*. Retrieved March 7, 2008 from <http://groups.ucanr.org/CCP/files/34394.pdf>
- Campbell, D., Glunt, E., Bockman, S., Sirotnik, B. W., & Little, J. (2003, March). *Evaluating the California community and faith-based initiative*. Paper presented at the 2003 Spring Research Forum of the Independent Sector, Bethesda, MD.
- Campbell, M. K., Denmak-Whnefried, W., Symons, M., Klasbeek, W. D., et al. (1999). Fruit and vegetable consumption and prevention of cancer: The Black Churches United for Better Health Project. *American Journal of Public Health*, 89(9), 1390–1396.
- Chaves, M. (2002). Religious organizations: Data resources and research opportunities. *American Behavioral Scientist*, 45(10), 1523–1549.
- Clerkin, R. M., & Gronbjerg, K. A. (2007). The capacities and challenges of faith-based human service organizations. *Public Administration Review*, 67(1), 115–126.
- Cnaan, R. A., & Boddie, S. C. (2001). Philadelphia census of congregations and their involvement in social services delivery. *Social Services Review*, 75(4), 559–580.
- Cnaan, R. A., & Boddie, S. C. (2006). Setting the context: Assessing the effectiveness of faith-based social services. *Journal of Religion & Spirituality in Social Work*, 25(3/4), 5–18.
- Compassion Capital Fund National Resource Center. (2005). *Measuring outcomes. Intermediary development series*. Washington, DC: Department of Health and Human Services. Retrieved from [http://www.acf.hhs.gov/programs/ccf/resources/gbk\\_pdf/om\\_gbk.pdf](http://www.acf.hhs.gov/programs/ccf/resources/gbk_pdf/om_gbk.pdf)
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Cook, T. D., & Campbell, D. T. (1979). In *Quasi-experimentation: Design and analysis issue for field settings*. Boston: Houghton Mifflin.
- Cordray, D. S., & Fischer, R. L. (1994). Synthesizing evaluation findings. In J. S. Wholey, H. P. Hatry, & K. E. Newcomer (Eds.), *Handbook of practical program evaluation* (pp. 198–231). San Francisco: Jossey-Bass.

- Deb, P., & Jones, D. (2003). Does faith work? A preliminary comparison of labor market outcomes of job training programs. In S. S. Kennedy & W. Bielefeld (Eds.), *Charitable Choice: First results from three states* (pp. 57–64), Indianapolis, IN: Center for Urban Policy and the Environment, School of Public and Environmental Affairs, Indiana University-Purdue University Indianapolis.
- DeHaven, M. J., Hunter, I. B., Wilder, L., Walton, J. W., & Berry, J. (2004). Health programs in faith-based organizations: Are they effective? *American Journal of Public Health, 94*, 1030–1036.
- DiIulio, J. J. (2000, Fall). Godly people in the public square. *The Public Interest*, 141.
- Ebaugh, H. R., Pipes, P. F., Chafetz, J. S., and Daniels, M. (2003). Where's the religion? Distinguishing faith-based from secular social service agencies. *Journal for the Scientific Study of Religion, 42*(3), 411–426.
- Erwin, D. O., Spatz, T. S., Stotts, R. C., & Hollenberg, J. A. (1999). Increasing mammography practice by African American women. *Cancer Practice, 7*(2), 78–85.
- Fagan, P. F., Horn, C., Edwards, C. W., Woods, K. M., & Caprara, C. (2007, March). *Outcome-based evaluation: Faith-based social service organizations and stewardship*. Washington, DC: The Heritage Foundation.
- Falcone, R. A., Brentley, A. L., Ricketts, C. D., Allen, S. E., & Garcia, V. F. (2006). Development, implementation and evaluation of a unique African-American faith-based approach to increase automobile restraint use. *Journal of the National Medical Association, 98*(8), 1335–1341.
- Farley, C. & Hackman, S. (2006). Ready4Work in brief. Interim Outcomes are in: Recidivism at half the national average. *In Brief, 4*. Philadelphia: Public/Private Ventures.
- Farley, C., & McClanahan, W. S. (2007). Ready4Work in brief. Update on outcomes; Reentry may be critical for states, cities. *In Brief, 4*. Philadelphia: Public/Private Ventures.
- Ferguson, K. M., Wu, Q., Spruijt-Metz, D., & Dyrness, G. (2007). Outcomes evaluation in faith-based social services: Are we evaluating faith accurately? *Research on Social Work Practice, 17*(2), 264–276.
- Fischer, R. L. (2001). The sea change in nonprofit human services: A critical assessment of outcomes measurement. *Families in Society, 82*, 561–568.
- Fischer, R. L. (2004). The devil is in the details: Implementing outcome measurement in faith-based organizations. *Nonprofit Management & Leadership, 15*(1), 25–40.
- Fischer, R. L., & Stelter, J. D. (2006). Testing faith: Improving evaluation research on faith-based human services. *Journal of Religion & Spirituality in Social Work, 25*(3/4), 105–122.

- Fink, B., & Branch, A. Y. (2005). *Promising practices for improving the capacity of faith and community-based organizations*. Prepared for Administration of Children and Families, U.S. Department of Health and Human Services. May.
- Florida Department of Corrections. (2000, September). *Comparing Tomoka Correctional Institution's faith-based dorm (Kairos Horizons) with non-participants*. Tallahassee, FL: Bureau of Research and Data Analysis.
- Fox, S. A., Stein, J.A., Gonzalez, R. E., Farrenkopf, M., & Dellinger, A. (1998). A trial to increase mammography utilization among Los Angeles Hispanic women. *Journal of Health for the Poor and Underserved, 9*, 309–321.
- Hangley, B., & McClanahan, W. S. (2002). *Mustering the armies of compassion in Philadelphia: An analysis of one year of literacy programming in faith-based institutions*. Philadelphia: Public/Private Ventures. Retrieved December 10, 2004, from <http://www.ppv.org>
- Hatry, H., Van Houten, T., Plantz, M., & Greenway, M. (1996). *Measuring program outcomes: A practical approach*. Alexandria, VA: United Way of America.
- Hodge, D. R., & Pittman, J. (2003). Faith-based drug and alcohol treatment providers: An exploratory study of Texan providers. *Journal of Social Service Research, 30*(1), 19–40.
- Independent Sector. (2003). *The role of faith-based organizations in the social welfare system—A report on the 2003 spring research forum*. Retrieved September 3, 2003, from <http://www.independentsector.org/PDFs/srf03report.pdf>
- Jeavons, T. H. (2004). Religious and faith based organizations: Do we know one when we see one? *Nonprofit and Voluntary Sector Quarterly, 33*, 140–145.
- Johnson, B. R. (2002). Assessing the impact of religious programs and prison on recidivism: An exploratory study. *Texas Journal of Corrections, 28*, 7–11.
- Johnson, B. R. (2004). Religious programs and recidivism among former inmates in prison fellowship programs: A long-term follow-up study. *Justice Quarterly, 21*(2), 329–354.
- Johnson, B. R., De Li, S., Larson, D. B., & McCullough, M. (2000). A systematic review of the religiosity and delinquency literature. *Journal of Contemporary Criminal Justice, 16*(1), 32–52.
- Johnson, B. R., & Larson, D. B. (2003, June). *The Innerchange Freedom Initiative: A preliminary evaluation for a faith-based prison program*. Report from the Center for Research on Religion and Urban Civil Society, University of Pennsylvania. Philadelphia: Center for Research on Religion and Urban Civil Society.
- Johnson, B.R. Larson, D.B. & Pitts, T.C. (1997). Religious programs, institutional adjustment, and recidivism among former inmates in prison fellowship programs. *Justice Quarterly, 14*(1), 145–166.

- Johnson, B. R., & Siegel, M. B. (2003). *The great escape—How religion alters the delinquent behavior of high-risk adolescents*. Philadelphia: University of Pennsylvania, Center for Research on Religion and Urban Civil Society.
- Johnson, B. R., Tompkins, R. B., & Webb, D. (2002). *Objective hope—Assessing the effectiveness of faith-based organizations: A review of the literature*. Philadelphia: University of Pennsylvania, Center for Research on Religion and Urban Civil Society.
- Johnson, C. L. (2000). *The Imani Project: Challenging missionary Baptist men to mentor adolescent African-American males*. Unpublished doctoral dissertation, Drew University.
- Jucovy, L. (2003). *Amachi: Mentoring children of prisoners in Philadelphia*. Philadelphia: Public/Private Ventures and Branch Associates. Retrieved December 10, 2004, from [http://www.ppv.org/ppv/publications/publications.asp?search\\_id=3](http://www.ppv.org/ppv/publications/publications.asp?search_id=3)
- Kennedy, S. S., & Bielefeld, W. (2003). *Charitable Choice: First results from three states*. Indianapolis, IN: Center for Urban Policy and the Environment.
- LaVigne, N. G., Brazzell, D., & Small, K. (2007). *Evaluation of Florida's faith- and character-based institutions. Final report*. Washington, DC: The Urban Institute.
- Leake, R., Green, S., Marquez, C., Vanderburg, J., Guillaume, S., & Gardner, V.A. (2007). Evaluating the capacity of faith-based programs in Colorado. *Research on Social Work Practice, 17*(2), 216–228.
- Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.
- MacMaster, S. A., Jones, J. L., Rasch, R. F. R., Crawford, S. L., Thompson, S., & Sanders II, E. C. (2007). Evaluation of a faith-based culturally relevant program for African-American substance users at risk for HIV in the southern United States. *Research on Social Work Practice, 17*(2), 229–238.
- Matthews, A. K., Berrios, N., Darnell, J. S., & Calhoun, E. (2006). A qualitative evaluation of a faith-based breast and cervical cancer screening intervention for African-American women. *Health Education & Behavior, 33*(5), 643–663.
- Mears, D. P., Roman, C. G., Wolff, A., & Buck, J. (2006). Faith-based efforts to improve prisoner reentry: Assessing the logic and evidence. *Journal of Criminal Justice, 34*, 351–367.
- Modesto, K. F. (2006). *Not by faith alone: A preliminary outcomes evaluation of a faith-based welfare-to-work program*. Doctoral dissertation submitted to the faculty of the University of North Carolina at Chapel Hill.
- Monsma, S. V., & Mounts, C. M. (2002). *Working faith: How religious organizations provide welfare-to-work services*. Philadelphia: University of Pennsylvania, Center for Research on Religion and Urban Civil Society.

- Monsma, S. V., & Soper, J. C. (2003, September). *What works: Comparing the effectiveness of welfare-to-work programs in Los Angeles*. Philadelphia, PA: Center for Research on Religion and Urban Civil Society.
- Monsma, S. V., & Soper, J. C. (2006). *Faith hope & jobs: Welfare-to-work in Los Angeles*. Washington, DC: Georgetown University Press.
- O'Connor, T., Su, Y., Ryan, P., Parikh, C., & Alexander, E. (1997). *Detroit transition of prisoners: Final evaluation report*. Washington, DC: Prison Fellowship Ministries.
- Powell, L. H., Shahabi, L., & Thoresen, C. E. (2003) Religion and spirituality: Linkages to physical health. *American Psychologist*, 58(1), 36–52.
- Ragan, M. (2004). *Faith-based vs. secular: Using administrative data to compare the performance of faith-affiliated and other social service providers*. The Roundtable on Religion and Social Welfare Policy, Retrieved May 9, 2005, from [http://www.religionandsocialpolicy.org/docs/research/Benchmarking\\_report\\_12-23-04.pdf](http://www.religionandsocialpolicy.org/docs/research/Benchmarking_report_12-23-04.pdf)
- Raudenbush, S. W. (1994). Random effects models. In H. Cooper & J. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 301–321). New York: Russell Sage Foundation.
- Regenerus, M. D. (2002). *Making the grade—The influence of religion upon the academic performance of youth in disadvantaged communities*. Philadelphia: University of Pennsylvania, Center for Research on Religion and Urban Civil Society.
- Rock, J. (2002, October). *Stepping out on faith: New York City's Charitable Choice demonstration program*. Albany, NY: SUNY, Rockefeller Institute of Government, The Roundtable on Religion and Social Welfare Policy.
- Roman, C. G., Wolff, A., Correa, V., & Buck, J. (2007). Assessing intermediate outcomes of a faith-based residential prisoner reentry program. *Research on Social Work Practice*, 17(2), 199–215.
- Rosenthal, R. (1979). The “file drawer problem” and tolerance for null results. *Psychological Bulletin*, 86, 638–641.
- Rosenthal, R. (1994). Parametric measures of effect size. In H. Cooper & J. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 231–244). New York: Russell Sage Foundation.
- Schorling, J. B., Roach, J., Siegel, M., Baturka, N., et al. (1997). A trial of church based smoking cessation interventions for rural African-Americans. *Preventive Medicine*, 26, 92–101.
- Scott, J. D., Montiel, L. M., Keyes-Williams, J., & Han, J. S. (2003, June). *The scope and scale of faith-based social services: A review of the research literature focusing on the activities of faith-based organizations in the delivery of social services* (2nd ed.). Albany, NY: SUNY, Rockefeller Institute of Government, The Roundtable on Religion and Social Welfare Policy.

- Sherman, A. L. (2006). Strengthening the street saints: Intermediaries providing capacity-building assistance to faith-based organizations. In R.V. Anglin (Ed.), *Building the organizations that build communities: Strengthening the capacity of faith- and community-based development organizations* (pp. 73–88). Washington, DC: Office of Policy Development and Research, U.S. Department of Housing and Urban Development.
- Sider, R. J., & Unruh, H. R. (2004). Typology of religious characteristics of social service and educational organizations and programs. *Nonprofit and Voluntary Sector Quarterly*, 33(1), 109–134.
- Singer, P., & Friel, B. (2007, January 5). Leaps of faith. *National Journal*. Retrieved February 28, 2008, from <http://nationaljournal.com/about/njweekly/stories/2007/0105nj1.htm>
- Smith, E. D. (1992). Hypertension management with church-based education: A pilot study. *Journal of the National Black Nurses Association*, 6, 19–28.
- Smith, S. R., Bartkowski, J. P., & Grettenberger, S. (2005). *Comparative views on the role and effect of faith in social services*. Albany, NY: SUNY, Albany Rockefeller Institute of Government, The Roundtable on Religion and Social Welfare Policy.
- Smith, S. R., & Sosin, M. R. (2001). The varieties of faith-related agencies. *Public Administration Review*, 61, 651–670.
- Stahler, G. J., Kirby, K. C., & Kerwin, M. E. (2007). A faith-based intervention for cocaine-dependent black women. *Journal of Psychoactive Drugs*, 39(2), 183–190.
- Twombly, E. C. (2002). Religious versus secular human service organizations: Implications for public policy. *Social Science Quarterly*, 83(4), 947–961.
- U.S. Department of Health and Human Services. (2002, June). The Compassion Capital Fund and the Faith- and Community-based Initiative, HHS fact sheet. Retrieved September 4, 2003, from <http://www.hhs.gov/news/press/2002pres/20020605.html>
- U.S. General Accounting Office. (2002). *Charitable Choice: Federal guidance on statutory provisions could improve consistency of implementation*. GAO-02-887. Washington, DC: author.
- U.S. Government Accountability Office. (2006). *Faith-Based and Community Initiative: Improvements in monitoring grantees and measuring performance could enhance accountability*. GAO-06-616. Washington, DC: author.
- Wall Street Journal*. (2003, June 20). Jesus saves: How President Bush found himself hugging a murderer in the White House. Editorial. Retrieved September 10, 2003, from <http://www.opinionjournal.com/taste/?id=110003652>.



- Wallace, J. M. (2002). *Is religion good for adolescent health? A national study of American high school seniors*. Philadelphia: University of Pennsylvania, Center for Research on Religion and Urban Civil Society.
- White House. (2008, February). *The Quiet Revolution—The President's Faith-Based and Community Initiative: A seven-year progress report*. Retrieved March 20, 2008, from <http://www.whitehouse.gov/government/fbci/reports.html>
- Wilson, R. J., Picheca, J. E., & Prinzo, M. (2005, May). *Circle of support and accountability: An evaluation of the pilot project in South-Central Ontario*. Correctional Service of Canada.
- Working Group on Human Needs and Faith-Based and Community Initiatives. (2002). *Finding common ground: 29 recommendations of the working group on human needs and faith-based and community initiatives*. Washington, DC: author.
- Working Group on Human Needs and Faith-Based and Community Initiatives. (2003). *Harnessing civic and faith-based power to fight poverty*. Washington, DC: author.