More of What?

Issues Raised by the Fort Bragg Study

John R. Weisz, Susan S. Han, and Sylvia M. Valeri
University of California, Los Angeles

The sobering findings of the Fort Bragg study illustrate why ambitious demonstration projects must be combined with objective outcome evaluations. The study does suggest that "more is not always better" (L. Bickman, 1996), but more of what? Little is known about the specific interventions that were combined to form the Fort Bragg system of care, so the study does not really reveal what failed or what needs to be changed. Moreover, there is no evidence that the specific treatments used had any empirical support. Combining and systematizing various treatments may not produce improved outcomes if the treatments are not effective in the first place. Costly demonstration programs that combine untested treatments may be a poor investment. A better strategy may be to develop and test an array of well-documented treatments for an array of child and family problems, creating the building blocks needed for effective systems of care in the future.

It is an intuitively appealing idea: Identify a set of mental health services for children and adolescents, organize the services into a coherent system or continuum of care, employ case managers to ensure that children and families get the services, and remove barriers to access (e.g., by providing transportation or taking therapy into the home). The result should be improved mental health for the youngsters thus served. But the findings of the Fort Bragg study (Bickman, 1996; Bickman et al., 1995) serve as a reminder that ideas may be both intuitively appealing and wrong. Mental health professionals and parents of troubled children, not to mention taxpayers, owe a good deal to the investigators and the framers of the Fort Bragg initiative (see especially Behar, 1996) who had the foresight to link this demonstration project to an independent evaluation. Arguably, the most important legacy of the program is the information generated by that evaluation. The report is a disciplined, dispassionate look at a particular model of intervention; its findings raise important issues about mental health care for children and families, aims and strategies for treatment development and outcome research, and the use of tax dollars to develop effective services.

Bickman's (1996) report indicated that the integrated continuum of care demonstration program developed at Fort Bragg had no better effect on short-term clinical or functional outcomes than did the traditional services available to children in the comparison sites. Numerous analyses (reported in Bickman, 1996; Bickman et al., 1995) have indicated that the absence of effects was not due to poor implementation of the continuum of care, faulty outcome assessment, or an equivalent intervention program at the comparison site. The seemingly robust null findings for child outcomes appear to be inconsistent with system of care theoretical models that stress the benefits of integrated services combined with case management. Even those who care little about theory might well ask how such an extensive and expensive program could have so little additional clinical impact. This important study raises both questions and concerns.

What Happened in the Interventions?

A particularly important question is what actually went on within the discrete services that were joined to form the Fort Bragg system? Like others invested in child mental health care, we would have preferred to see positive outcomes of the system intervention at the end of this massive venture. However, given the limited documentation of specific intervention processes, it could be argued that neither positive, negative, nor null outcomes could be fully informative. Certainly it is important to know that the program did not produce the intended effects, and Bickman and colleagues (Bickman, 1996; Bickman et al., 1995) have done well in identifying service categories, documenting how personnel spent their time, surveying clinical records, and providing other information on the structure and configuration of services included in the program. In fact, the Fort Bragg study was designed to test a system, not individual services. Thus, the investigators deserve special credit for the information they did provide on component services. However, the information available to them, and thus to read-

---

John R. Weisz, Susan S. Han, and Sylvia M. Valeri, Department of Psychology, University of California, Los Angeles.

Some of the research discussed in this article was supported by National Institute of Mental Health (NIMH) Grants R01 MH38450 and R01 MH38240. The preparation of this article was supported by Research Scientist Award K05 MH01161 and Research Grant R01 MH49522 from NIMH.

Correspondence concerning this article should be addressed to John R. Weisz, Department of Psychology, Franz Hall, University of California, Los Angeles, 405 Hilgard Avenue, Los Angeles, CA 90095-1563.

Vol. 52, No. 5, 541–545

ers, reveals relatively little about the specific intervention procedures used by the intervenors in their work with children and families, the work that constituted the heart of the various services.

Because the providers involved were apparently not relying on empirically tested interventions guided by manuals, the study could not provide the detailed documentation of treatment content and sequence to which readers of the clinical trials literature have become accustomed. And because there is no well-developed method for documenting treatment procedures and processes in conventional clinical practice, there is no process-related information to substitute for manual descriptions. Complicating the picture is the likelihood that treatments in the same nominal category (e.g., outpatient psychotherapy) are apt to have been different from one therapist–child–parent combination to the next; in principle, clinically, this is appropriate and desirable, but it underscores the need for documentation of the processes involved, case by case.

Because such documentation is lacking, it is difficult to know exactly what has been shown not to work. This complicates the task of planning further steps in research on child treatment and systems of care. In the absence of content and process information, findings regarding treatment outcomes, either positive or negative, have limited value. When treatment works well, the specific processes involved need to be identified so they can be repeated, refined, and ultimately disseminated. And when treatment leads to disappointing results, it is important to identify the specific processes involved so that these can be stopped and alternative approaches tried. The Fort Bragg report indicated that little significant benefit was derived from integration of an array of services, even though the integration involved increased costs. Thus, it could be argued that further attempts to integrate those particular services may be nonproductive and perhaps wasteful. But what were those particular services? The study does not answer this question at a level of specificity that would enable treatment planners to reconfigure the system of care with better interventions, appropriately retrain the providers, and test a genuine alternate approach.

The absence of information about treatment content and process also limits the use of subgroup analysis for heuristic purposes. In principle, hypotheses about promising treatment approaches can be generated by identifying best-outcome and worst-outcome subsets of a full treatment sample and then backtracking to group differences in intervention specifics. In the case of Fort Bragg, too little is known about intervention specifics for this strategy to be very productive, at least with regard to treatment content and process (although the general strategy may still be useful for other purposes).

This is a criticism of the field and its methodological limitations, not of the Fort Bragg study in particular. The Fort Bragg findings are useful in a number of important ways, but the project also underscores the need for better ways of assessing treatment content and process in the child area (see, e.g., Shirk & Russell, 1996). The need is particularly profound in research involving the nonmanualized interventions that are most characteristic of clinical practice. Without good measures of therapy content and process in clinical practice, effectiveness studies may lack the precision needed to describe, instruct, and direct. Instead, findings may indicate that ''services'' or ''systems'' either worked or didn't work, but without a clear enough picture of the component procedures and processes to shape the next steps.

Were the Individual Services Effective?

Information on treatment process, of course, derives value through companion information on treatment outcome, and herein lies another issue raised by the Fort Bragg study. Among the possible reasons why integration of services produced so little effect is that the individual services that were integrated were not very effective. Recent work (reviewed in Weisz, Donenberg, Han, & Kauneckis, 1995; Weisz, Donenberg, Han, & Weiss, 1995; Weisz & Weiss, 1993; Weisz, Weiss, & Donenberg, 1992) has brought both good and bad news about child mental health interventions. The good news is that most controlled treatment outcome studies in the clinical trials literature reveal respectable positive effects of child interventions (see Casey & Berman, 1985; Kazdin, Bass, Ayers, & Rodgers, 1990; Weisz, Weiss, Alicke, & Klotz, 1987; Weisz, Weiss, Han, Granger, & Morton, 1995). The bad news is that the modest number of studies testing typical interventions for referred children in clinical settings reveal small to negligible effects (see Kutash & Rivera, 1996; Rivera & Kutash, 1994; Weisz, Donenberg, Han, & Kauneckis, 1995; Weisz, Donenberg, Han, & Weiss, 1995).

Taken together, the findings suggest that the beneficial effects of child interventions that have been demonstrated in clinical trials research may not be replicated in the real-world clinical settings where most child treatment takes place. This may not be particularly surprising given the fact that the child treatments most widely practiced in clinical settings are generally not derived from the pool of empirically supported therapies and in fact may frequently differ markedly from those therapies (see, e.g., Kazdin et al., 1990; Weisz & Weiss, 1993). There appears to be a broad gap between the world of treatment outcome research, where systematic treatments are devised and tested empirically, and the world of clinical practice, where treatments often evolve independently of the empirical literature (see L. Cohen, 1979; L. Cohen, Sargent, & Sechrest, 1986; Morrow-Bradley & Elliott, 1986; Raw, 1993). In summary, what may be happening is that conventional clinical practice frequently involves treatments that are not empirically derived or empirically supported and perhaps also not particularly effective (further tests of this possibility are needed, as discussed in Weisz, Donenberg, Han, & Weiss, 1995). The possibility bears scrutiny in light of the current debate over the use of empirically supported treatments in clinical psychology (see American Psychological Association, Task Force on Promotion and Dissemination of Psychological Procedures of Clinical Psychology, 1995; Chambless, 1996, in press; Kovacs, 1995).

As for the Fort Bragg program, there is no indication that the various treatments used therein were selected because they had prior empirical support. From a strictly empirical perspective, then, there may be little basis for assuming that the interventions were, in fact, effective (see Bickman, 1996, p. 699). And, of course, combining and systematizing treatments may not improve outcomes if the treatments are not effective in the first place. In other words, it is possible that the interventions brought together in the Fort Bragg program had such modest individual effects that there was little to be gained by multiplying and coordinating them.[1]

Developing Effective Treatments: A Vote for Small Science

The Fort Bragg findings suggest an important message for treatment and services researchers in an era of limits and downsizing. Although a logical case can be made for a top-down strategy, testing large child mental health programs before examining their component effects, it is not clear that this strategy has improved mental health outcomes for children. The problem is that these large multicomponent programs have so often produced minimal benefit (see examples below). This repeated pattern suggests that the cart may have been placed before the horse in research on child mental health care. Before mounting tests of the impact of integrated services, perhaps researchers should assess the singular impact of those discrete services they plan to integrate. If the individual building blocks are found to be weak, then it may be premature to combine them into a larger structure.

More broadly, perhaps efforts to understand and improve child and family mental health care could be strengthened by a shift in emphasis from big science to little science. The $80 million spent on the Fort Bragg program could have funded a very large number of focused studies on specific interventions for specific child and family problems. The track record of such focused interventions, summarizing more than 300 outcome studies in the child area (see Weisz, Donenberg, Han, & Weiss, 1995), is strong, with an average unadjusted effect size of 0.77, approaching the J. Cohen (1988)-derived standard for a ''large'' effect. If more support were provided for the development of such specific interventions, the result might be a rich and ever-expanding array of empirically supported treatments for an ever-expanding range of child and family problems. As this array of treatments is built, treatment planners, clinicians, and researchers would be in a stronger position to create systems of care that integrate treatments of known benefit. Research dollars may be best spent, at present, on development of the building blocks needed for successful integration of services in the future.

The fact that the Fort Bragg evaluation showed disappointing results does not mean, of course, that no integration of component treatments can work. Programs that combine previously distinct, empirically supported interventions may work quite well. As an example, Henggeler and colleagues (e.g., Henggeler et al., in press; Henggeler, Schoenwald, & Pickrel, 1995) have shown very positive effects of their multisystemic therapy program for serious juvenile offenders. The program combines various empirically supported interventions and delivers them in home, school, and other real-world settings. The well-replicated findings of the Henggeler group illustrate that linking discrete interventions into larger programs may be a very good idea, provided that the discrete interventions have been shown to be beneficial in their own right.

Need for User-Friendly, Clinic-Ready, Empirically Supported Treatments

Of course, development of beneficial treatments in controlled research would not necessarily mean that such treatments would—or even could—be used within systems of care. As suggested above, the gap between the outcome research community and the clinical practice community is broad and deep, such that most interventions developed by outcome researchers languish in their laboratories, largely unused in clinical practice. In noting this possibility, we do not attribute sole (or even primary) responsibility to clinical practitioners. Treatment outcome researchers have not been particularly effective at making their work accessible to practitioners or at facilitating incorporation of their treatments into clinical practice. The gap between outcome research and clinical practice can be bridged by the development and adaptation of efficacious treatments for use with seriously disturbed children and families in real-world clinical settings. The experience and judgment of experienced clinicians could be invaluable in these efforts. It is important to learn what adjustments must be made in empirically derived treatments to make those treatments usable and effective with children like those in the Fort Bragg project. If empirically supported individual interventions are ever to be successfully integrated into community systems of care, those interventions must be user-friendly and effective in the real world of clinical practice.

Demonstrations Need Outcome Evaluations

Although development and refinement of focused treatments require increased attention, there can be little doubt that big science will continue in some form. This being the case, another implication of the Fort Bragg findings bears attention: Demonstration programs designed to help children and families need to be subjected to inde-

---

[1] As an alternative interpretation, it is possible that the individual treatments integrated at Fort Bragg and the treatments being used in the comparison sites were all highly effective. If this were the case, then outcomes of the unintegrated services might have been so good that the integrated services in Fort Bragg simply could not outstrip this ceiling effect. The absence of a no-treatment control group in the Fort Bragg project ruled out a test of this possibility, but the evidence we are aware of on conventional mental health services for children does not support this interpretation.

pendent outcome evaluation to determine whether children and families have actually been helped. Some might argue that this would turn demonstration projects into experiments, but this is precisely the point. Demonstration projects that are not also experiments may tell little more than the following two self-evident facts: (a) Programs can be developed, and (b) money can be spent. What most taxpayers actually want to know is whether developing the programs and spending the money lead to beneficial effects.

Imagine, for a moment, what the situation might be had there been no evaluation of the Fort Bragg project. One possibility is that system of care funding would continue to flow to this site into the indefinite future, at a cost of $17 million per year plus inflation, and it is conceivable that other sites would have been added. After all, consumer satisfaction with Fort Bragg was quite impressive (see below; see also Heflinger, Sonnichsen, & Brannan, 1996). Even if the project had been halted after a few years, in the absence of an evaluation component, relatively little would have been learned for all the expense. Although it would have been established that a system of care program could be organized and services could be delivered, little would be known about whether taking these steps actually improves outcomes for children and families. Thanks to the framers of this project, there is now evidence that systematizing services in this way in this setting did not significantly enhance outcomes. This information, though discouraging, is certainly useful.

These null results are consistent with findings from a number of other large-scale intervention and prevention programs. For example, a study of North Carolina's $25 million per year "Willie M. Program" for violent and assaultive youth showed little evidence of positive effects (Weisz, Walter, Weiss, Fernandez, & Mikow, 1990). And a recent evaluation of the nation's most widely used school-based substance use prevention program, the $750 million per year Drug Abuse Resistance Education (DARE) Program, showed little evidence that DARE reduces drug use (Ennett, Tobler, Ringwalt, & Flewelling, 1994; Ringwalt et al., 1994). The evidence from these and other evaluations (see, e.g., Burns, Farmer, Angold, Costello, & Behar, in press; Goldman, Morrissey, & Ridgely, 1994; Shern, Wilson, & Coen, 1994) suggests that investing large sums of money in intervention programs is no guarantee that the programs will be effective. This, in turn, argues for objective, independent outcome evaluations of all such interventions.

Back to Basics in Outcome Research

Finally, the Fort Bragg project underscores the value of a disciplined adherence, where feasible, to certain basic principles of outcome research. One such principle is that mental health outcomes should be assessed, in part, using clinical measures of established validity. Although such currently popular measures as consumer satisfaction ratings are useful for certain purposes, it is important to distinguish clearly between such ratings and measures of clinical outcome. Bickman and colleagues (Bickman, 1996; Heflinger et al., 1996) found high levels of overall consumer satisfaction with outpatient treatment, and notable subgroup differences as well, with significantly higher ratings from Fort Bragg parents than from comparison parents on all nine of their questions about outpatient treatment. But when the focus shifted to more standard clinical measures of child behavioral and emotional problems, psychopathology, and impairment of functioning, the apparent benefit of the Fort Bragg system of care evaporated (see Bickman et al., 1995; Lambert & Guthrie, 1996). In this case, sole reliance on the consumer satisfaction ratings could have led to an incorrect conclusion about the clinical impact of treatment.

Other findings of the study serve as reminders that clinical improvement following treatment may or may not reflect true intervention effects; control or comparison groups, or comparison conditions (in single-group or single-case designs), are required for meaningful interpretation of change over time. Children in the Fort Bragg program showed substantial clinical improvement coincident with therapy. The Fort Bragg youngsters improved markedly on clinical measures over the initial 6-month lag and continued to improve over the 1-year lag, with good 12-month rates on Jacobson and Truax's (1991) reliable change index (see Lambert & Guthrie, 1996). In the absence of a comparison group, such findings might have seemed to support the effectiveness of the system of care program. However, the fact that a matched comparison group, with no system of care, showed slightly higher rates of reliable clinical change over the same period places change in the Fort Bragg program in its proper interpretive context.

This finding echoes a trend seen in other clinical research in both the effectiveness (see Weisz & Weiss, 1989) and efficacy (Weiss & Weisz, 1990) traditions: Individuals who are referred to clinics and are judged appropriate candidates for treatment, but who do not actually receive treatment, tend to show significant clinical improvement over time nonetheless. Among the possible explanations of this trend, one must consider the nature and timing of treatment seeking. People tend to seek treatment when problems and distress are at unusually high levels. Thus, scores and ratings on measures of psychological dysfunction are apt to be at an atypically high level at the time treatment is sought (Time 1 in the typical outcome study). On average, weeks or months later, those scores and ratings are likely to have regressed toward the person's mean level of functioning, regardless of the form of treatment or whether treatment of any kind actually took place. Whether this hypothesis or some alternative explains the tendency of comparison–control groups to improve over time, the tendency itself underscores the risks inherent in drawing conclusions about treatment effectiveness based on change over time in treated individuals alone.

Concluding Thoughts

It is, of course, disappointing to discover that a major intervention has not produced the desired effects. But a good deal can be learned from this frank and careful evaluation of the Fort Bragg program. The findings highlight the need for methodological advances, and they raise important issues about approaches to treatment development, the relation between research and practice, and how to balance big and little science for maximum benefit to children and families. Thus, in a number of ways, the findings of the Fort Bragg study hasten the day when effective treatments can be used to create truly effective systems of care.

REFERENCES

American Psychological Association, Task Force on Promotion and Dissemination of Psychological Procedures of Clinical Psychology. (1995). Training in and dissemination of empirically-validated psychological treatments: Report and recommendations. *The Clinical Psychologist, 48,* 3–23.

Behar, L. (1996). Policy implications of the evaluation of the Fort Bragg Child and Adolescent Mental Health Demonstration Project. *The Journal of Mental Health Administration, 23,* 118–121.

Bickman, L. (1996). A continuum of care: More is not always better. *American Psychologist, 51,* 689–701.

Bickman, L., Guthrie, P. R., Foster, E. M., Lambert, E. W., Summerfelt, W. T., Breda, C. S., & Heflinger, C. A. (1995). *Evaluating managed mental health services: The Fort Bragg experiment.* New York: Plenum Press.

Burns, B. J., Farmer, E. M. Z., Angold, A., Costello, E. J., & Behar, L. (in press). A randomized trial of case management for youths with serious emotional disturbance. *Journal of Clinical Child Psychology.*

Casey, R. J., & Berman, J. S. (1985). The outcome of psychotherapy with children. *Psychological Bulletin, 98,* 388–400.

Chambless, D. L. (1996, June). Identification of empirically supported psychological interventions. *Clinician's Research Digest, 14*(Suppl. Bulletin 14), 1–2.

Chambless, D. L. (in press). In defense of dissemination of empirically supported psychological interventions. *Clinical Psychology: Science and Practice.*

Cohen, J. (1988). *Statistical power analyses for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cohen, L. (1979). The research readership and information source reliance of clinical psychologists. *Professional Psychology, 10,* 780–785.

Cohen, L., Sargent, M., & Sechrest, L. (1986). Use of psychotherapy research by professional psychologists. *American Psychologist, 41,* 198–206.

Ennett, S. T., Tobler, N. S., Ringwalt, C. L., & Flewelling, R. L. (1994). How effective is drug abuse resistance education? A meta-analysis of Project DARE outcome evaluations. *American Journal of Public Health, 84,* 1394–1401.

Goldman, H. H., Morrissey, J. P., & Ridgely, S. M. (1994). Evaluating the Robert Wood Johnson Foundation program on chronic mental illness. *Millbank Quarterly, 72,* 37–48.

Heflinger, C. A., Sonnichsen, S. E., & Brannan, A. M. (1996). Parent satisfaction with children's mental health services in a children's mental health managed care demonstration. *The Journal of Mental Health Administration, 23,* 69–79.

Henggeler, S. W., Borduin, C. M., Schoenwald, S. K., Pickrel, S. G., Rowland, M. D., & Cunningham, P. B. (in press). *Multisystemic treatment of antisocial behavior in youth.* New York: Guilford Press.

Henggeler, S. W., Schoenwald, S. K., & Pickrel, S. G. (1995). Multisystemic therapy: Bridging the gap between university- and comunity-based treatment. *Journal of Consulting and Clinical Psychology, 63,* 709–717.

Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59,* 12–19.

Kazdin, A. E., Bass, D., Ayers, W. A., & Rodgers, A. (1990). Empirical and clinical focus of child and adolescent psychotherapy research. *Journal of Consulting and Clinical Psychology, 58,* 729–740.

Kovacs, A. L. (1995). "We have met the enemy and he is us!" *The Independent Practitioner, 15,* 135–137.

Kutash, K., & Rivera, V. R. (1996). *What works in children's mental health services? Uncovering answers to critical questions.* Baltimore: Brooks.

Lambert, E. W., & Guthrie, P. R. (1996). Clinical outcomes of a children's mental health managed care demonstration. *The Journal of Mental Health Administration, 23,* 51–68.

Morrow-Bradley, C., & Elliott, R. (1986). Utilization of psychotherapy research by practicing psychotherapists. *American Psychologist, 41,* 188–197.

Raw, S. D. (1993, March). Does psychotherapy research teach us anything about psychotherapy? *The Behavior Therapist,* pp. 75–76.

Ringwalt, C. L., Greene, J. M., Ennett, S. T., Iachan, R., Clayton, R. R., & Leukefeld, C. G. (1994). *Past and future directions of the D.A.R.E. Program: An evaluation review.* Research Triangle Park, NC: Research Triangle Institute.

Rivera, V. R., & Kutash, K. (1994). *Components of a system of care: What does the research say?* Tampa: University of Florida, Center for Children's Mental Health, Research and Training, Florida Mental Health Institute.

Shern, D. L., Wilson, N. R., & Coen, A. S. (1994). Client outcomes II: Longitudinal client data from the Colorado Treatment Outcome Study. *Millbank Quarterly, 72,* 123–148.

Shirk, S., & Russell, R. L. (1996). *Change processes in child psychotherapy: Revitalizing treatment and research.* New York: Guilford Press.

Weiss, B., & Weisz, J. R. (1990). The impact of methodological factors on child psychotherapy outcome research: A meta-analysis for researchers. *Journal of Abnormal Child Psychology, 18,* 639–670.

Weisz, J. R., Donenberg, G. R., Han, S. S., & Kauneckis, D. (1995). Child and adolescent psychotherapy outcomes in experiments versus clinics: Why the disparity? *Journal of Abnormal Child Psychology, 23,* 83–106.

Weisz, J. R., Donenberg, G. R., Han, S. S., & Weiss, B. (1995). Bridging the gap between laboratory and clinic in child and adolescent psychotherapy. *Journal of Consulting and Clinical Psychology, 63,* 688–701.

Weisz, J. R., Walter, B. R., Weiss, B., Fernandez, G. A., & Mikow, V. A. (1990). Arrests among emotionally disturbed violent and assaultive individuals following minimal versus lengthy intervention through North Carolina's Willie M. Program. *Journal of Consulting and Clinical Psychology, 58,* 720–728.

Weisz, J. R., & Weiss, B. (1989). Assessing the effects of clinic-based psychotherapy with children and adolescents. *Journal of Consulting and Clinical Psychology, 57,* 741–746.

Weisz, J. R., & Weiss, B. (1993). *Effects of psychotherapy with children and adolescents.* New York: Sage.

Weisz, J. R., Weiss, B., Alicke, M. D., & Klotz, M. L. (1987). Effectiveness of psychotherapy with children and adolescents: A meta-analysis for clinicians. *Journal of Consulting and Clinical Psychology, 55,* 542–549.

Weisz, J. R., Weiss, B., & Donenberg, G. R. (1992). The lab versus the clinic: Effects of child and adolescent psychotherapy. *American Psychologist, 47,* 1578–1585.

Weisz, J. R., Weiss, B., Han, S., Granger, D. A., & Morton, T. (1995). Effects of psychotherapy with children and adolescents revisited: A meta-analysis of treatment outcome studies. *Psychological Bulletin, 117,* 450–468.