Like so many others, I was drawn to psychotherapy as a profession out of my personal experiences with suffering and the help I experienced from therapy. Adolescence and early adulthood had been extremely difficult for me. Progress out of the quagmire in which I was stuck appeared to derive from my extensive experiences as a therapy patient. My exploration of my internal life in therapy converged with an intellectual interest in psychoanalytic theory arising, initially, from my readings in the Freudian left. Social thought and depth psychology converged in psychoanalysis. When the opportunity arose for psychoanalytic training, I grabbed it, even

---

I thank Peter Kreiner and Nina Kammerer for comments on drafts of this chapter. The evaluation of Project Second Beginning described herein was supported by a grant from the Substance Abuse and Mental Health Administration’s Center for Substance Abuse Prevention to Associates for Human Potential. The research at the Boston Institute for Psychotherapy was generously supported by the institute out of very limited funds. These sources of support are gratefully acknowledged. The views expressed here are solely the views of the author and are not necessarily those of any organization.

223
as I obtained the more traditional credentials: a master's in counseling and a doctorate in clinical psychology.

Unlike many who are drawn to psychoanalytic theory and practice, I had an extensive background in the “hard sciences,” having studied advanced mathematics and related fields for many years. In addition, I am temperamentally a skeptic. Taking anything on faith is contrary to my nature. I therefore quickly wondered about the evidential bases for various therapeutic propositions.

My skepticism received a boost from the experience of case conferences in clinics where I worked. Traditionally, a case conference involved a patient being presented to, and frequently interviewed by, a prominent therapist from outside the agency. What I quickly noticed was that the guest expert never seemed to have trouble determining in half an hour exactly what ailed the patient and what should be done about it. For those of us in the audience, however, there was frequently a problem. Our local in-house expert usually also had determined the nature of the patient’s problems, and his—it always was a he in those days—prescription often differed radically from that of the outside expert. What were we to do? Here we had two distinct assessments and treatment prescriptions, each asserted with great confidence by experts we respected, but the prescriptions differed from each other. In many cases, if one was right, the other would clearly have to be wrong. If so, what happened to the hapless patients who went to the “wrong” therapist? How could they possibly get better? Furthermore, I knew that many patients didn’t go to therapists trained in the intricacies of psychoanalysis. In fact, in those days, many patients went to Gestalt therapists, encounter groups, client-centered therapists, and so forth. Although none of my personal acquaintances went to behaviorists, I knew that many patients did go in that direction. What happened to all these patients? Did they get better despite the limitations (as perceived by me) of the theoretical orientations of their therapists?

As a neophyte therapist, how was I to choose between the numerous rival claimants to my attention? Having a scientific background, my natural inclination was toward research. “Surely,” I thought, “the data will tell us which treatment approaches work.” Unfortunately, the available evidence is confusing at best. Luborsky, Singer, and Luborsky (1975) had summarized the available outcome data and came to the conclusion that “everyone has won and all must have prizes,” as the Dodo bird expressed it in Alice in Wonderland. All therapies appeared to work essentially equivalently well. Smith, Glass, and Miller (1980) came to much the same conclusion. More recently, Project Match, a $28 million attempt to match types of patients to alcoholism treatment, experienced a spectacular failure: No differences were found.

What differences that have been found between treatments have tended to favor behavioral treatments. To some degree, these studies could
be disturbing to a psychodynamically oriented therapist. But in these studies, researchers have tended to compare brief therapies for very circumscribed conditions, using outcome measures tailored to the particular targeted behavior problems. It took no great leap to see these studies as largely irrelevant to much of psychodynamic therapy as practiced with patients with multiple, often rather amorphous problems, such as being chronically unhappy and dissatisfied with one's life, having difficulties experiencing intimacy with others, and so forth. Many of the patients I saw had already completed one or more courses of focused brief therapy but often felt that their core problems had not even been identified, much less addressed, in the process.

For example, one patient I am currently seeing, Mr. King, came into treatment because of a depression that was interfering with his work functioning. He refused to even consider medication, which was just as well, because his depression remitted fairly quickly. He then was left with the feeling that he hadn't been very happy, at least since his mother died in adolescence. He had never had a satisfying relationship with a woman, and people in general were experienced as more of a bother and a threat to his stability than as pleasurable. Eventually he came to identify his central dilemma as having to decide whether he would want relationships with others. Perhaps life as a loner would be good enough for him and better than the suffering that involved attempting to interact with others. What could the clinical trials literature tell me about the best way to treat this patient? Nothing, as far as I could see.

One approach to this perceived irrelevance of outcome research to many psychodynamic practitioners involves the development of outcome measures that more faithfully measure the problems experienced by these types of patients. Such efforts are ongoing, but have proven difficult. The development of valid, useful psychological measurement techniques is normally hard, and measurement of fairly obscure constructs regarding aspects of psychological life that are not directly observable is extremely difficult. Psychotherapy researchers have, in general, found that patients' perceptions of their changes owing to therapy bear only a limited relationship to the changes perceived by the patients' therapists, even when the same construct is being assessed. If other observers or interviewers are brought in, yet more perspectives are generated. Thus, sophisticated psychotherapy research projects have taken to using multiple measures from multiple perspectives. The result is somewhat greater fidelity to clinical reality accompanied by much greater complexity of interpretation. In any case, studies using elaborate outcome packages haven't led to clearer guidelines for clinicians.

Although I have not given Mr. King a packet of standardized outcome

---

1Patient names cited here have been fictionalized.
measures, I suspect that he would exhibit moderate depression on a symptom questionnaire and might even exhibit difficulties with intimacy on one of my favorite self-report questionnaires, the Inventory of Interpersonal Problems (Horowitz, Rosenberg, Bauer, Ureno, & Villasenor, 1988; Soldz, Budman, Demby, & Merry, 1995b). However, no questionnaire or standardized interview that I am aware of would elucidate the problem he came to identify: whether to want relationships with others. The standardized questionnaires and interviews traditionally used in psychotherapy research, although providing much useful information about this man, would fail to inform the user of the nature of his problem from his perspective and whether that problem had improved.

One solution to the outcome problem is to develop methods of measuring individualized problems. Such methods have been regularly pursued. In fact, the one instrument I did administer to Mr. King was a form of George Kelly’s (1955) repertory grid. The grid was enlightening to me. It showed that Mr. King bifurcated people into two classes, those who were “rational” but “boring,” and those who were “irrational,” “emotional,” and incomprehensible to him. The use of the grid allowed us to refine our goals for treatment, but such subtleties would be missed in most psychotherapy research programs. And even if Mr. King’s problems could be measured adequately, I’m still faced with deciding how to treat him in the absence of much guidance from the treatment literature.

GUIDANCE FROM PROCESS RESEARCH

The solution to this lack of clear guidance for clinical practice from much of psychotherapy outcome research that I believed in a decade ago, and still do believe in to some extent, involves the need to develop a clearer understanding of the curative mechanisms of psychotherapies. One way of elucidating these mechanisms is to study the process of therapy. Over the decades, much research has taken this tack (Soldz, 1990). The most notable results from this work so far have involved the therapeutic or working alliance between patients and therapists (Horvath & Greenberg, 1994). Many studies have found that the quality of the alliance predicts outcomes, as traditionally measured, in many therapies, including psychodynamic, cognitive, and group (Budman, Soldz, Demby, Feldstein, & Springer, 1989). The alliance literature is important in indicating, as psychodynamic therapists have long believed, that the relationship between the patient and therapist is important to the patient’s outcome.

An implication of this work on the alliance is that a therapist should strive to maximize the alliance with a patient. Or is it? Unfortunately, so far we have no studies examining whether deliberate attempts to improve the alliance lead to improved outcomes. In particular, we need to know
whether there are techniques that can be applied that will lead to improved alliances—and, possibly, outcomes—for the more severely disturbed personality-disordered patients who are extremely difficult to treat and who frequently drop out prematurely or have poor outcomes (Strupp, 1980).

Another general conclusion from the psychotherapy process literature is that patients who are more disturbed with poor histories of relationships with others are not successfully treated by most therapists. This conclusion was driven home to me when I was writing a survey article on psychodynamic process research for a clinical audience (Soldz, 1990). I started out summarizing a number of classic studies in the area. However, I did not start out with a conclusion in mind. When I got to the section of the paper where I needed to write about the implications of the research, I was struck by the consistency of findings: “Difficult” patients with poor histories of relationships were rarely successfully treated. Strupp (1980) expressed the results of detailed analyses of successful and unsuccessful cases in one of his classic psychotherapy research studies: “In our study we failed to encounter a single instance in which a difficult patient’s hostility and negativism were successfully confronted or resolved” (p. 954). Because these patients typically form poor alliances with their therapists, we need to know whether these alliances can be improved, and whether improving the alliances will lead to better retention and outcome. Unfortunately, the process research literature so far provides little guidance regarding these types of questions. This research didn’t help me decide which, if any, of the impressive “experts” I heard at case conferences I should trust.

In general, my impression of the process literature is that it provides an exciting first step in helping us understand the change processes in psychotherapy, but that most of this work is far from providing concrete clinical guidance. I have hopes that this research will evolve in a more applicable direction, but I am far from convinced. If this work is to be applied, one of two things will have to happen. Either research should focus on the ability to deliberately change therapeutic processes, as in the example of alliance research where investigation would focus on the ability to improve alliances and determine whether improved alliances lead to improved outcomes, or researchers could explicate in fine detail processes that lead to patient improvement. This direction was strongly promoted about a decade ago when Rice and Greenberg (1984) published an important book titled Patterns of Change. Many of us expected this book and its strong reception among psychotherapy researchers to spur process research in this direction. If researchers could describe change processes in detail, there would be a reasonable chance that such knowledge could be incorporated into therapists’ techniques. Just as therapists now learn from hearing how a respected colleague would treat a case, they could learn from research that indicated that which styles on responding, in what particular circumstances, with what types of patients, would be most helpful.
For whatever reasons, however, the expected surge in research on change processes did not occur. I expect that the reasons this research has not surged are twofold. One set of reasons has to do with the extreme difficulty of getting such research funded. In the past decade, the federal funding sources that provide almost all funding for large-scale psychosocial treatment research have focused most of their resources on randomized clinical trials that compare treatments, based on the pharmacological model for comparing two drugs. Process research is virtually unfundable at this point. Thus, only small-scale studies, based on volunteer or graduate student labor can be completed. Although every year yields quite a number of interesting studies of this type, they often remain too small to lead to cumulative knowledge that could be usefully applied by clinicians.

A second set of reasons that this type of process research hasn't flourished involves the lack of good, precise, theories of human change processes that can serve as the basis for this type of fine-grained analysis. In the absence of these detailed theories, researchers are left adrift when confronting the mass of material generated by even a brief therapeutic interaction.

As I treat Mr. King, I am aware of little process research that can aid me. Knowing that he has a history of very poor relationships with others, beginning in childhood, the process research would alert me to the dangers that he would terminate prematurely. I did not need research to raise my awareness in this area, however. Mr. King was referred to me by a colleague whom he had seen for a few sessions and promptly discarded. Our sessions were full of intense interactions in which I would say something that would lead to him feeling misunderstood and thinking of terminating. A large part of our interactions were spent repairing these ruptures in our alliance. Although it is nice to know that research supports the importance of repairing ruptured alliances (Safron, Muran, & Samstag, 1994), I cannot honestly say that my attempts in this area were improved through my knowledge of this research.

GROUP THERAPY RESEARCH

My own experiences with process research primarily took place at the Mental Health Research Department of Harvard Community Health Plan (HCHP), where I was affiliated for the first 8 years of my research career. At HCHP, we investigated interactional group therapy. The first study, which I entered as the project was in progress, involved assessing the cohesiveness of brief (15-session) therapy groups. Our main finding was that the cohesiveness of groups, as rated by clinical observers of videotapes, was strongly related to the outcomes of the patients in the groups. As our group cohesiveness measure was developed to be a group analog of the therapeutic
alliance, we thus extended the work on the alliance in individual therapy to the group setting. I well remember the excitement when we ran our first data analyses and realized that the results were going to come out as predicted (Budman et al., 1989).

Yet, were there clinical implications of this finding? We weren’t sure. As we examined videotapes of low- and high-cohesiveness groups, it was clear that the groups differed in their functioning. It was also clear that the low-cohesiveness groups differed in composition of the patients, as well. The patients in them seemed to be less psychologically minded, among other things; they didn’t seem to have a particularly rich set of constructs for thinking and talking about the internal lives of themselves and the other group members. So, was low cohesiveness a cause of patient change, or simply a result of patient characteristics that prevented the patients from creating a group that could lead to positive change? We studied cohesiveness for a number of years (Budman, Soldz, Demby, Davis, & Merry, 1993), but never really answered this question.

Similar issues bedeviled our research into the behaviors of individual patients in the groups (Soldz, Budman, Davis, & Demby, 1993; Soldz, Budman, & Demby, 1992; Soldz, Budman, Demby, & Feldstein, 1990; Soldz, Budman, Demby, & Merry, 1995a). We found out a number of interesting things. For example, the patients who spoke the most in our groups were the most disturbed, and what they spoke about so much was themselves, rather than the other group members. Although this finding may not seem surprising when stated, many clinicians to whom I have described the study predict that the opposite will be the case until hearing the results; then they find these results obvious. We also found that personality-disordered group therapy patients could agree with each other and their therapists regarding the personalities of their fellow patients, but had little agreement with the others regarding their own personalities. We found, further, that the patients with the greatest agreement with fellow patients about their own personalities were those with the greatest degree of personality disturbance: This latter finding is never predicted by people to whom I describe the study.

As we studied group process and patient change in these groups, we gradually realized that we were hoping that patient personality could be changed by way of group therapy. Many of the patients who enter psychotherapy come not to have particular symptoms relieved, but because their personality patterns lead to problematic and unsatisfying interactions with the important other people in their lives. Could group therapy change these patterns? We created a set of 18-month groups for patients with personality disorders, based on interpersonal principles and informed by our research experiences (Budman, Cooley, et al., 1996). These groups were modestly successful: Those patients who stayed in them tended to get better, but many of the patients, especially those with borderline
personality diagnoses, did not complete the groups (Budman, Demby, Soldz, & Merry, 1996). This latter result is reminiscent of Strupp's (1980) comments above about the failures of the therapists in his study to successfully treat hostile patients with poor histories of relationships.

In our years studying therapy groups, we made many discoveries, but their meaning was never unambiguous enough for me—the inveterate skeptic—to feel that we knew enough to issue treatment recommendations. In my opinion, it would have taken several research careers and an amount of grant money that was unimaginable even in less fiscally conservative times to have advanced our knowledge enough to have given us a shot at making definitive recommendations. Other members of the research group were less reticent about making recommendations based on limited data. In either case, there did not appear to be any large audience of clinicians or policy makers interested in listening to any recommendations we would have made.

EVALUATION

If randomized clinical trials and process research lead to little knowledge of use to practicing clinicians, does this mean that research cannot usefully aid practice? Many have come to this conclusion. By accident, I was exposed to an area where research and practice intersect that provides a useful model for future efforts.

A colleague who had to leave town due to death in the family called me and asked if I would write an evaluation plan for a substance abuse prevention grant. I am now embarrassed to admit that my first question to him was "What is an evaluation?" In my training and experience in research, I had learned much about statistics, research design, and the relationship of theory and research, but the existence of program evaluation was never mentioned. What I discovered as I helped this colleague was that evaluation was an established domain where research skills were used to provide data-based guidance to program administrators and policy makers.

A couple of years later, after a few rewrites and resubmissions, the project for which I wrote that first evaluation plan was funded by the Federal Center for Substance Abuse Prevention (CSAP). This project, named Project Second Beginning (PSB), involved the provision of case management to substance abusing women who were either parents or pregnant. As the project evolved, a treatment model was formed that was based on the relational model of women's development that had been developed by theorists and researchers at the Stone Center at Wellesley College. This model involved a clinically trained staff (psychologists, social workers, and
nurses) providing what came to be called relational case management to the
women (Markoff & Cawley, 1996). Relational case management involved
the clinician forming a supportive relationship with the client that provided
a model of healthy gratifying relationships while simultaneously helping her
navigate through the morass of social service agencies in order to meet her
life needs for food, shelter, day care, vocational training, and the like.

When I started out as project evaluator, I naively assumed that my
job was simply to assess whether the program was successful in helping the
women. I quickly realized that I was expected to play many roles. While I
was to report on the clients' outcomes, I was also expected to provide both
qualitative and quantitative information on the functioning of the project.
This information, referred to in the evaluation literature as process evalu-
ation data (not to be confused with psychotherapy process research), was
expected to be used both for documentation purposes and to improve the
project. The model for the relationship between the project and its eval-
uation that was frequently referred to by the CSAP staff and their con-
suitants involved the creation of a self-correcting system. The idea was
that evaluation data would be fed back to the project folks, who would
then use it to improve the project, in a never-ending cycle, at least until
our funding ran out.

Shortly after PSB was funded, I met with the newly hired program
staff to modify the evaluation plan I had quickly drawn up for the grant
in order to reflect the evolving program model. Without knowing how
controversial it was (Patton, 1986; Shadish, Cook, & Leviton, 1991), I
naturally developed a collaborative model whereby the evaluation of the
project was the creation and responsibility of all of us—administrators,
clinicians, and myself as the evaluator. We held numerous discussions re-
garding the staff's conceptualizations of the project, focusing on the
changes they expected to occur in their clients. We then jointly discussed
various instruments for measuring these changes, with several staff members
and I sharing responsibility for trips to the library to find relevant research
literature, and for phone calls to other projects that had faced similar eval-
uation issues so that we could learn from their experiences. Although I
contributed my research expertise as we selected instruments and designed
an evaluation protocol, there was no assumption that I was the only one
whose expertise was important.

What I did contribute was based on my years of experience in psy-
chotherapy research. As we selected outcome measures, I was aware of the
complexities of the measurement of change, including such issues as the
desirability of measuring change from multiple perspectives (such as self,
clinician, others); the value of including multiple measures to compensate
for the peculiarities of particular instruments; the importance of including
more than two time points when measuring change; and the danger of
assuming that an instrument necessarily measures what it purports to measure without clear validity evidence.

Nonetheless, our evaluation design was less than optimal due to the extremely limited resources at our disposal, as well as to our collective naivety. We assumed, for example, that our first experiences with clients who tended to stay a long time would generalize, leading us to schedule our follow-up measurements at too long an interval, which we gradually discovered was after many clients already had left. We were also well aware that we would ultimately not be able to definitively attribute to PSB any changes in the women since we were unable to include any control or comparison group. Nonetheless, we hoped that the evaluation data would provide at least some useful information regarding what happened to the women in the project.

After the project and evaluation had been going for a while, we had the pleasant experience of being called to task by our CSAP project officer for basing our project on the Relational Model of Women's Development and not measuring the relationship between the woman client and the case manager. In response, we decided to introduce into our evaluation materials the Working Alliance Inventory (Horvath, 1994), a measure of the alliance developed for psychotherapy research.

As the data were collected and analyzed, we developed preliminary findings. In addition to the bureaucratic requirement that results be communicated to CSAP periodically, we held meetings of all staff to discuss the emerging findings. I presented aspects of the data that I found either intriguing or confusing, leading to long discussions regarding possible meanings. Among the results we discussed collaboratively were those directly relevant to assessing how well the program was doing as well as those confusing relations between variables that I didn’t understand.

One set of results we often discussed concerned client retention. Overall, clients stayed a median of 7 months. This length of stay was less than we had originally planned, but was quite respectable when compared with other substance abuse projects. However, we also found that clients with more serious psychiatric symptomatology stayed significantly less long. The clinical staff had not been aware of this latter fact until I presented the data. A number of discussions ensued over the succeeding months regarding the reasons for this finding and ways of trying to identify and improve the treatment of these women. The working alliance appeared to predict length of stay, but, alas, the women with more psychiatric symptomatology had worse alliances. Despite these repeated discussions and the attempts of staff to be more attuned to the needs of these women, it appeared that we never successfully improved the treatment of this more disturbed group to the extent that they stayed as long as others.

Note the similarity to the findings from psychotherapy research discussed above.
In conducting the process evaluation, we were interested in innovative ways to document the treatment model developed by the PSB case managers. From our perspective, the primary function of this documentation was to communicate the cumulative knowledge being developed by the project. We were especially interested in finding ways to capture some of the implicit knowledge gained by clinicians in the course of their work. We experimented with two techniques for capturing this knowledge. One technique, concept mapping (Trochim, 1989; Trochim, Cook, & Setze, 1994), involves a structured brainstorming session, followed by computer manipulations, leading to a group interpretation of the resulting graphical output. Our general feeling was that this process did not lead to any deeper understanding.

We then decided to explore CareMaps, a way of systematically recording, in a matrix format, the procedures involved in a treatment process, based on the best practices known to a given group of clinicians. After the CareMap matrix is created, it can be used to document the treatment course of a given client. An essential aspect of the use of CareMaps is that variation from the expected course of treatment is systematically recorded and analyzed by way of variance analysis, either leading to improved protocols or to the development of separate CareMaps for identified subgroups. Second Beginning staff spent many long hours developing and refining our CareMap. The resulting product contained, in my opinion, an enormous amount of information about the steps and processes that could aid the substance abusing mothers targeted by the project. The clinical staff were less convinced of the value of their product. It seemed to me that their skepticism was due to the fact that the resultant CareMap represented knowledge that had become second nature to them through years of hard work, but that could be extremely valuable to others beginning such a project. Furthermore, it is possible that if the CareMap had been used systematically variance analysis would have helped improve treatment for certain clients, including, perhaps, those with psychiatric impairment.

As I was involved in evaluating PSB and a number of other demonstration projects, I witnessed a marked change in the relationship between research and practice among many substance abuse programs. The federal requirement for an evaluation of each project forced researchers, program administrators, and direct service staff to work closely together. Rather than familiarity breeding contempt, in many instances it bred increased respect for the perspectives and expertise of the others. Clinical staff and administrators were exposed to the processes and the insights of research and researchers at close range, leading to greater appreciation. At the same time, researchers got to know and respect the difficulties faced

---

3The CareMap name and concept are copyrighted by the Center for Case Management in South Natick, MA. I thank Kathleen Andolina of the center for her wonderful assistance during the development of the CareMap.
by treatment programs and the dedication of their staff. In the best instances, the creation of exemplary programs, and of excellent evaluations became a shared responsibility.

Furthermore, as programs sought this federal money, there was an increasing demand that any proposal be based in the existing "knowledge base" in the area, making it routine for those writing grants to examine existing research and evaluation literature as they designed a program. Initially, this examination was solely for the purpose of justifying a pre-existing idea, but gradually, the literature review became a source of information on what was known to work or not work. The grant writers and program administrators also became aware that the existence of data suggesting that their program was effective materially increased the odds of getting funded. As the director of one program in the retardation area to whom I had consulted on the design of information systems put it: "When we apply for funding, we have the data to show them exactly what their money will buy. As a result, we always get the money" (Joan Beadsley, personal communication, 1994). The tie between knowledge of relevant research and the existence of program data and the awarding of funding increased program administrators' respect for the research process. My sense was that these processes radically changed the relationship between research and practice in the substance abuse area, making them, in many cases, natural allies rather than wary adversaries. The federal demonstration grant program was thus, perhaps inadvertently, a demonstration to the field that research and practice can cooperate productively. Similar experiences are currently being forced on the mental health field by the cry for outcomes data from managed care companies and public funders.

BOSTON INSTITUTE FOR PSYCHOTHERAPY

Other important experiences that helped develop my thinking about the research-practice interface came out of my efforts in the past decade to develop a psychotherapy research program in an institute committed to training and practice in psychodynamic psychotherapy, the Boston Institute for Psychotherapy (BIP). Through a fortuitous accident, I connected with the BIP just as a couple of Institute graduates were completing their doctorates and were attempting to set up postdoctoral fellowships with a significant research component. They, along with myself and a BIP staff member who had a long history of social psychological research, decided to establish a program of research into the treatment provided in the BIP's Treatment Center. After a period of preparation, we decided to conduct regular outcome measurements on willing patients, while recording some therapies for future analysis of the patient-therapist interaction. Despite the expressed support of both the BIP board and program directors (the
senior management of the Institute), we encountered fierce resistance from a number of staff (Mordecai, Soldz, & Gumpert, 1993). Many clinicians appeared afraid that the research would be used to judge the quality of their work, while others were worried about its intrusive effects on the therapy they provided. There was also a small but vocal group who argued that traditional research into dynamic psychotherapy was useless at best, as the essence of this type of therapy involves subtle processes that can never be measured.

In order to confront these resistances, we sought to involve the BIP clinicians in the design of our research through conducting focus groups with them regarding the nature of outcomes in dynamic psychotherapy and through regular presentations to the treatment center's Case Conference. We were also instrumental in the establishment of a monthly Grand Rounds series that combined clinical and clinically relevant research presentations. For a number of years, I also taught the course on Freud and classical psychoanalysis to the second-year trainees in which I insisted on integrating research on psychoanalytic concepts with the historical and clinical material. The result of these efforts was a greater appreciation among many BIP staff of the value of research and attitudes ranging from tolerance to excitement about our efforts.

After several years of functioning, the research group became familiar with the CareMap concept and decided to experiment with its application to open-ended dynamic psychotherapy. In order to create a CareMap for a given group of patients, it is necessary to identify a relatively homogeneous subgroup of patients who can normally be expected to have a similar course of treatment. We decided to concentrate on young adults in their twenties who had identity and relationship difficulties because this type of patient was common in the BIP treatment center. Although we were impressed with our ability to identify expectable events for this population—such as "She will become concerned about dependency on the therapist within the first 6 months of treatment"—we also became aware that we did not know enough about the normative course of treatment to construct a CareMap. For example, our clinical experience led us to believe that some patients changed in a slow, continuous manner; some changed rapidly for short periods with long plateau periods of no clear change in between; and yet others got worse at certain points in the treatment, only to improve later. But were our clinical intuitions consonant with reality? We modified our goal of constructing a CareMap to one of exploring the nature of change over the course of treatment, at least as that change is viewed by the therapist.

We thus undertook the construction of an instrument to be completed by therapists on a monthly basis that would assess patient change on a broad range of dimensions relevant to psychodynamic therapy. Over a year later, our product, the Boston Institute for Psychotherapy Change
Assessment Tool (BIP-CAT) emerged. The BIP-CAT contains scales measuring patient functioning and change in four broad areas: problems and symptom, external life, internal life, and the therapeutic relationship.

The Problems and Symptoms section captures information about the problems and symptoms that are identified by either the patient or therapist. Given our orientation toward using the BIP-CAT to record the therapist’s perceptions of changes in the patient, the therapist is allowed to identify a problem or symptom that has not been explicitly identified by the patient.

The subsections of the other three sections, with sample items are as follows:

1. External Life
   Romantic relationships (Sample item: "Quality of sex life"); Friendships (Sample item: "Ability to see other’s point of view"); Family of origin relationships; Self-care (Sample item: "Pleasurable activities [Fun]"); Work (Sample item: "Presence of stable career goals")

2. Internal Life
   Self (Sample item: "Sense of well-being"); Object relations (Sample item: "Significant distortion of others"); Defenses; Thought (Sample item: "Awareness of wishes/longings"); Affect

3. Therapeutic Relationship
   Content of sessions (Sample item: "Pt notes patterns in feelings, thoughts, and actions"); Frame (Sample item: "Therapist finds it difficult to negotiate frame"); Transference (Sample item: "Pt idealizes therapist"); Countertransference (Sample item: "Therapist experiences boredom"); Intervention types

Each item on the BIP-CAT is rated at intake and every 6 months thereafter on its severity as a problem for the patient. In addition—and this is the heart of the instrument—it is rated every month in terms of whether the patient changed (either improved or got worse) on the item. If completed systematically, the BIP-CAT would provide an "X-ray" of the therapy, allowing us to view its vicissitudes over time.

One of the unique features of the BIP-CAT is that therapists are asked to rate both patient level of functioning and change over the past month for the scales in the first three areas. It is our hope that the direct ratings of change will allow us to detect small changes that would need to accumulate over many months before being assessable as change of even one point on a direct rating of patient functioning.

As we completed the BIP-CAT and wrote a detailed manual, we started presenting it to the BIP administrators and clinicians. Many of them responded to it with intense excitement, because they could see the direct links between the clinical constructs important to them and the dimen-
sions of the instrument. Thus, they did not feel that we were trying to fit
the square peg of psychodynamic therapy into the round hole solely of
symptom relief. They could see that through the BIP-CAT, we were at-
ttempting to measure such subtle clinical constructs as object relations or
defense use. As the therapists found that they could relate to the BIP-CAT,
they also became more interested in our other research projects. Mean-
while, the BIP directors group, of which I was a member as the director of
research, became more interested in using data on our patients in order to
manage the Institute. Where did the patients come from? Did certain types
of patients drop out “prematurely”? Would certain modifications of intake
procedures lead to greater patient retention? The answers to these questions
clearly require the development of data systems. The BIP is thus moving
toward the development of evaluation systems, with the assistance of the
research group. These modifications are leading the BIP toward including
research and evaluation in the routine functioning of the Institute, thus
breaking down barriers between research and clinical practice.

CONCLUSION

The experiences I have described, as well as many others, have led
me to rethink the relationship of research to practice. Initially, I was hope-
ful that research could provide a guide to practice with individual patients,
that it could help me decide between the advice of the various case con-
ference presenters. I no longer think that the primary impact of research
is likely to be at this level. However, there are a number of other ways in
which research can affect practice more directly. Such efforts will involve
collaborative efforts between clinicians and researchers, with neither side
having primacy. To successfully impact practice, these collaborative efforts
will require the creation of a strong, vibrant, researcher-practitioner alli-
ance.

The primary lesson derived from my experiences over the years is that
bridges are built between research and practice primarily when researchers
and practitioners work together. This working together may be driven by
internal forces, as has occurred at the BIP, or by external forces, as with
the Demonstration Project programs. But, as in so many other areas of life,
mutual understanding and respect occur when people of different back-
grounds find themselves working together on common projects of mutual
interest.

Thus, I and my fellow clinicians may not treat Mr. K. and other
patients with a treatment directly based on research. Yet, if a collaborative

*This issue had already been investigated by Boston Institute for Psychotherapy staff and
reported in a journal article (Pollack, Mordecai, & Gumpert, 1992).
researcher–practitioner alliance can be built, the treatment of future Mr. King may be improved through the feedback and research-based input that can be glimpsed in some of the experiences I have described. Psychotherapy may never be purely an applied science, as some researchers dream. With luck, however, it may become more suffused with the insights and knowledge of science than is now the case. If that future should develop, it is my belief and desire that researchers, therapists, and their patients will all find psychotherapy to be a richer, more valuable experience.

REFERENCES


