Progression of Therapy Research and Clinical Application of Treatment Require Better Understanding of the Change Process

Alan E. Kazdin, Yale University

The stage model of therapy research focuses on the development of treatment from pilot work, through randomized controlled clinical trials, to tests in clinic settings. A goal of the model is to develop effective treatments that can be used clinically. The present comments begin with a similar goal but emphasize the importance of a broader agenda designed to understand therapy. A central thesis is that developing effective treatments depends heavily on investigations that address critical scientific questions; particularly, what are the mechanisms through which therapy operates and under what conditions is therapy likely to be effective and why? The comments argue for a portfolio of research that addresses a broader range of questions and encompasses more diverse methods of evaluating treatment. Breadth and diversity are not ends in themselves but will be essential to obtain the requisite knowledge to effect optimal changes in clinical applications of treatment.

Key words: **psychotherapy research, therapy evaluation.** ***[Clin Psychol Sci Prac 8:143–151, 2001]***

Psychotherapy research has advanced considerably in the past few decades, as reflected by the number of empirical investigations, the development and identification of empirically supported treatments, and the methodological standards of individual investigations (Kendall & Chambless, 1998; Nathan & Gorman, 1998; Snyder & Ingram, 2000). Even so, there are good reasons to systematize therapy research and to foster progress in treatment development. If one considers the flow of therapy research and development, presumably critical points in the pipeline include developing a treatment, determining its effectiveness, and bringing the treatment to market (i.e., service delivery). Quite clearly, the pipeline is clogged in at least a couple of places. At the innovation and development phase, there is a glut of treatments. Several hundred therapy techniques have been developed and are in use clinically. The majority of these have never been studied empirically. Hence, the treatments do not move farther down the pipeline. This is no deterrent to further technique innovation and development. New treatments are constantly emerging and make the task of providing a fixed count impossible (see Kazdin, 2000). Even though "new" sometimes is the eye of the beholder, this is somewhat moot in the larger scheme of things. We are not at a loss for developing treatments.

A later segment in the pipeline is also clogged. There are several treatments that have been designated as empirically supported in light of the rigorous evidence they have in their behalf (Kendall & Chambless, 1998; Lonigan & Elbert, 1998; Nathan & Gorman, 1998). Yet the treatments are not brought to market in a way that they can be used in service delivery settings. One of the questions that keeps these treatments clogged in the pipeline is whether they are effective when extended to clinical settings. Empirically supported treatments are viewed by many as ready to move to the next stage, but the requisite evidence in clinical settings, leaving aside the daunting dissemination task, may impede this step.

There seems to be no systematic or well-articulated way of moving from clinical innovation to a well-developed treatment that has impact on clients in clinical

Address correspondence to Alan E. Kazdin, Department of Psychology, Yale University, P.O. Box 208205, New Haven, CT 06520-8205.

© 2001 AMERICAN PSYCHOLOGICAL ASSOCIATION D12

settings. Perhaps it would be easy to ignore the geological pace in progressing from therapy research and development to widespread implementation if it were not for the seemingly parallel sequence of drug research and the movement from laboratory research on mechanisms to clinical trials and dissemination of medications to clinical practice. The path of drug research is intricate, complex, outside of public view, and is strewn with serendipity, failed leads, and not-quite effective treatments (see Stone & Darlington, 2000). Even so, what is in public view is the conspicuous unveiling of new treatments with palpable effects for serious mental and physical disorders. Perhaps if we as clinical investigators organized therapy research a bit better and provided investigators with a sequence of questions, stages, or methods of evaluating treatment, the movement to well-developed and useful treatments would be more rapid and evident.

Rounsaville, Caroll, and Onken (this issue; see also Onken, Blaine, & Battjes, 1997) have outlined a sequence of studies to move psychotherapy from clinical development to application and to ensure that the requisite empirical evidence is obtained along the way. A stage model of therapy research is proposed that emphasizes the methodological requirements of studies, as progress is made from piloting treatment, conducting controlled clinical trials, and ensuring that treatment can be extended or transported to clinical settings.[1] The model is a systematic way to redress the pipeline issues I have noted and to provide a template that could exert significant impact on therapy research.

The proposed stage model can be examined from at least two perspectives, depending on the primary impetus leading to its development. First, the model can be examined from the standpoint of delineating the research agenda or priorities of a funding agency or institution. In this context, the issues pertain to the goals of the agency, the type of research the agency wishes to foster, the types of grant proposals legitimate for consideration, and the essential components of these proposals. Second, the model can be examined from the standpoint of the broader scientific agenda of therapy research. In this context, the issues pertain to the key goals and questions of therapy research and more generally to how one proceeds from development of treatment to clinical application. The funding agency and scientific agenda are not completely independent, but the distinction is meaningful and frames the context for these comments.

Much of the proposed stage model focuses on the priorities and interests of funding agencies and grant applications. This is evident from the discussion of the difficulties in obtaining funds for pilot work, the requisite information for a grant proposal designed to fund pilot work, the criteria for judging grant proposals, the time frame for completing pilot work in relation to the demands of funding cycles, and other issues. It is difficult, and perhaps even inappropriate, to evaluate the stage model in relation to the priorities and interests of one or more funding agencies. Agencies are free to delineate a research agenda, to convey the means of pursuing that agenda, and to direct potential applicants to the types of research they will fund. The information from the stage model will be quite helpful to potential grant applicants and may promote further thought on the entire process of what to fund to optimize the eventual yield from therapy research. Also, the model takes a formal position on the utility of pilot work and tests of generality of treatment effects. Pilot work and tests of generality of therapy occasionally are viewed as conceptually bereft because they address crassly empirical questions and rarely progress from description to explanation. The model is intended to legitimize both types of studies (stages I and III) and show how they relate to randomized controlled clinical trials (RCTs), which are universally appreciated.

Apart from funding issues, it is meaningful to evaluate the stage model in light of broader treatment goals—that is, the extent to which the model addresses critical scientific questions about therapy. Does the model help us move to where we wish to go in developing the knowledge base and, if so, how? Consideration of the broader scientific questions shows, I believe, that the model is silent on many critical issues that are central to understanding therapy. Moreover, a key goal of the model—identifying treatments that can be used effectively in clinical settings—may not be optimally served by the sequence of studies that the stage model proposes. Here I examine the model as well as the larger goals toward which it is directed.

WHAT DO WE WANT TO KNOW ABOUT TREATMENT?

Before considering the stages of research and the types of studies that are needed, we ought to ask, what do we want to know about therapy (i.e., what are the goals)? Only then is it meaningful to ask what type of research is needed to obtain the answers (i.e., what are the means).

Stated generally, it seems that we want to know if a treatment is effective, why and how it achieves change, what the conditions are that contribute to its effectiveness, and ideally why and how these conditions contribute to change. We also want to know if treatment can be extended to clinical practice and if it is effective in such extensions. The practical agenda makes the scientific agenda so critically important, a point worth elaborating.

Why Treatment Works

Elaborating the underpinnings of treatment is absolutely critical in the progression of therapy research. Understanding therapy begins with some level of theory to explain the change process and then direct tests of the hypothesized mechanisms or process. There are all sorts of armchair explanations of why this or that therapy produces change. These are rarely tested, and studies are rarely even designed in ways in which mechanisms that cause change can be inferred from the findings (see Kazdin, 2000).

The paucity of empirical work on the reasons that therapies are effective is unfortunate. First, as mentioned already, there are several hundred therapy techniques in use, leaving aside eclectic hybrids and treatment combinations. Perhaps there is a small set of change mechanisms or processes that span several techniques. Theory and direct tests of these processes might bring order and parsimony to the current status of multiple interventions.

Second, the effects of psychotherapy are broad. Variations of psychotherapy improve mental health (e.g., reduce suicidality, depression), ameliorate physical conditions (e.g., pain, high blood pressure, recovery from surgery or illness), and, indeed, affect life (e.g., increases in fertility) and death (e.g., increases in survival among terminally ill patients) (see Kazdin, 2000). How any or all of these effects might be achieved begs for theory and supportive research.

Third, there are an unlimited number of potential moderators of therapeutic change (i.e., variables on which the effectiveness of a given treatment may depend). For example, characteristics of the client (e.g., family history, severity of dysfunction, comorbidity, and personality attributes), the context in which the client functions (e.g., family life, current stressors), characteristics of the therapist (e.g., experience, personality style), and, of course, characteristics of treatment administration, all might influence outcome as main effects or part of interactions. With an unlimited number of moderators from which to draw, theory can focus empirical tests by posing those factors likely to be significant and why.

Fourth, an obvious goal of treatment is to optimize therapeutic change. Theory can help us understand the processes that account for therapeutic change and hence those processes that ought to be fostered and maximized. Without theory, the facets of therapy that can be used to improve treatment outcome will be difficult to identify and investigate. Also, without the resulting knowledge of critical processes, treatment manuals are likely to include all sorts of ingredients that make little difference and underemphasize those that do.

Fifth, theory of therapeutic change is important because of its broader relation to psychological science. There are many processes in everyday life that contribute to adjustment and adaptive functioning. Examples include participating in religion, chatting with friends, exercising, undergoing hypnosis, and writing about sources of stress. Therapy research is not only about treatment techniques but also about a broader question; namely, how does one intervene to change biological, social, emotional, and behavioral functioning? Theories that elaborate how therapy works might have generality for understanding human functioning more generally and vice versa.

Conceptual views are needed about what treatment is designed to accomplish and through what processes or mechanisms. The guiding question is how does this treatment achieve change? The answer may involve basic processes at different levels (e.g., neurotransmitters, stress hormones, memory, learning, information processing, motivation). Theories of change must be followed by direct empirical tests. Do the intervention techniques, methods, and procedures within treatment sessions actually affect those processes that are considered to be critical to the treatment model? For example, it may be that changes in cognitions are critical among depressed patients receiving cognitive therapy. If so, the pertinent cognitions ought to change during treatment and there ought to be a special relation between change in cognitive processes and client improvement, beyond a mere correlation between change in cognitions and change in symptoms. At least three steps are required to conduct the requisite research: specifying a conceptual view of the processes or factors responsible for change, developing measures of these processes, and showing that these processes change during therapy and before therapeutic

change. This latter requirement is needed to establish the time line (i.e., processes are changing and are not merely concomitant effects of symptom improvement). The reasons that treatments may work have been amply discussed but they are rarely carefully studied.

Identifying Conditions That Influence Treatment Effects

Tests of the conditions on which effective application of treatment depends are critical. What are the conditions for effective application of treatment or the variables (moderators) that influence effectiveness? The effectiveness of treatment can depend on all sorts of factors, as already mentioned. Conceptualization and empirical findings regarding the clinical problem can inform the search for moderators. For example, we know that many sexually abused children are likely to develop cognitions that the world is a dangerous place, that adults cannot be trusted, and that one's own efforts to influence the world are not likely to be effective (Wolfe, 1999). Based on this understanding of the problem, one might predict that sexually abused youths with these cognitions would respond less well to treatment, as measured by posttreatment prosocial functioning. If these cognitions are not altered in treatment, the children may be restricted in social activities compared to similar children without these cognitions. Perhaps another study using this information would evaluate if the effectiveness of treatment could be enhanced by including a component that focuses on these cognitions. In general, the search for moderators ought to be guided by our understanding of the factors related to dysfunction or to the change process. The search also may be guided by clinical experience, which is often an excellent place to begin. The task is to end up with a conceptual view of how, why, and for whom therapy is effective, along with supportive evidence, but we can begin just with good ideas.

There is a danger in merely testing for moderating effects without ever moving to understanding how they operate. The difference is between descriptive and explanatory research (i.e., studies that show a relation and studies that explain the basis of that relation). For example, one might find gender differences in response to treatment. This is not the end but rather begins the process of trying to understand what it is about gender that relates to therapeutic change. Gender may be a proxy for other factors (e.g., cognitive development or patterns of social interaction), and once these factors are considered, gender differences may disappear or become nugatory. Understanding why or how the moderator works can be important for both theoretical and applied reasons.

Tests of Generalization and Applicability

As treatment is shown to produce change in a particular context or setting, it is valuable to evaluate the generality of the findings across other dimensions and domains. Can this treatment be applied clinically, and are the effects similar to those obtained in research? Tests of generality of a treatment are subsumed in the prior comments about testing moderators of treatment, insofar as moderators refer directly to the conditions that influence outcome. It is worth treating tests of generality separately. Ideally, tests of treatment moderators such as client or therapist characteristics draw on theory or prior research to predict the factors on which treatment effects may depend. Tests of generality are less conceptually inspired and more application oriented (i.e., can treatment be applied in different ways, to different people, and in different settings). Tests of generality may require several factors to vary from the laboratory conditions in which treatment was developed (e.g., dose and variation of treatment, characteristics of the clients, therapists, and settings), and these are not easily separated in an analytic way that permits delineation of their contribution to the outcome.

Clearly, we want to know if treatment is transportable from laboratory to clinical settings. This keen interest has led to a call for more tests of treatments in clinical settings. Perhaps paradoxically, the best investment for generality of treatment effects may be further laboratory studies of treatment. The problem with laboratory tests (so-called efficacy studies) has been that they have not systematically developed the knowledge base. Typically, studies focus on outcomes in controlled settings, which is clearly important, but no more than an initial step to the critical questions. Additional work is needed that focuses on understanding the underpinnings of treatment and on how to optimize therapeutic change. We simply do not know what is needed to make treatment effective or optimally effective and how change comes about. Treatments as currently tested include a package of ingredients and the interventions that, if tested in clinic settings, would be examined without knowing the active and critical components.

General Comments

The stage model appears to be guided by the efficacy/effectiveness issue (Hoagwood, Hibbs, Brent, & Jensen, 1995). Efficacy refers to the outcomes of treatment in the context of well-controlled laboratory settings, and effectiveness refers to outcomes in the context of clinical and more real-life applications. In the stage model, these concerns are reflected in stage II and stage III, respectively. Tests of generality of treatment seem very much like apple pie and mother(parent)hood. Who could object? As tests of generality are currently formulated, I believe most people ought to object. The reason is that without knowing why treatment works (mechanisms) and for whom treatment works (moderators), extensions to clinical settings are not likely to work very well. We will not know what the conditions for optimally effective application are and what components in a treatment manual are important, necessary, and facilitative.

There has been a stream of articles on manualization of treatment, including discussions of "to manualize" or "not to manualize," fine-grained distinctions of how to manualize, and, of course, a flood of books designed as manuals. Scant attention is devoted to the fact that ingredients in the manual for most treatments have hardly been shown to be central at all. An amazing array of components is included in most manuals that reflect practices and recommendations that make "good clinical sense" or are based on "clinical experience." The clinical scientist knows that these terms are used when there are no data; also, the terms usually appear right before the sentence that reads, "more research is needed."

All recommendations in a given manual cannot be evaluated empirically. This is why the mechanisms of therapeutic change are so essential. What are the facets of treatment that lead to change, and what psychological, social, and biological processes do they activate that promote change? A host of recommendations in manuals and little bits of superstitious behavior here and there would be less problematic if we knew the kernel of what must be included and how to activate the processes known to be responsible for change.

Knowing how to optimize the effects of treatment is particularly important in extensions of treatment to clinical settings. There are multiple factors of clinical settings that can dilute treatment effects, such as reduced monitoring of therapist adherence to treatment protocols, as compared with tests in laboratory settings, increased diversity of clinical patients to whom treatment is applied, and some restrictions on treatment delivery (e.g., more sporadic sessions because of checkered patient attendance). It is all the more important to know the critical mechanisms of treatment and how to optimize their effects. This does not mean one never tests a treatment in clinical settings without knowing why it works; it does mean that the neglect of why questions is short-sighted.

I have not represented well the complexity of the task of therapy research. First, there is much we want to know about treatment. Among the key questions to guide treatment research, the following are salient:

1. What is the impact of treatment relative to no-treatment?
2. What components contribute to change?
3. What treatments can be added (combined treatments) to optimize change?
4. What parameters can be varied to influence (improve) outcome?
5. How effective is this treatment relative to other treatments for this problem?
6. What patient, therapist, treatment, and contextual factors influence (moderate) outcome?
7. What processes within or during treatment influence (mediate) outcome?
8. To what extent are treatment effects generalizable across problem areas, settings, and other domains?

There is a range of answers to these questions, as reflected in the domain (and measures) that are used to evaluate outcome. The effects of treatment usually are evaluated on presenting symptoms (e.g., anxiety), but there are additional domains, such as impairment, positive adaptive functioning, and quality of life, that may be equally or more pertinent in relation to short- and long-term functioning. The difficulty is that the verdict about treatment (i.e., whether it is effective or has clinically significant impact) can depend on the measurement domain. Two treatments that are equally effective in altering presenting symptoms can vary in their impact on other domains (e.g., family functioning) (Szapocznik et al., 1989). Thus, conclusions about treatment very much depend on the outcome domain.

I have not mentioned the time line of therapy outcome

evaluation. Conclusions about the efficacy of a treatment or relative efficacy of different treatments may vary greatly depending on when assessments are conducted. All sorts of combinations and permutations have been found. Thus, sometimes treatments that differ from each other or from controls at posttreatment do not differ at follow-up or vice versa; sometimes changes over the course of treatment are small and nonsignificant but increase over time or vice versa (e.g., Kolvin et al., 1981; Meyers, Graves, Whelan, & Barclay, 1996; Newman, Consoli, & Taylor, 1997). Clearly, conclusions about the effects of a given treatment relative to a control condition or another treatment may vary at posttreatment and follow-up.

The complexities of therapy research are not minor annoyances, but rather central to the subject matter. Thus a model of research and interest in identifying clinically useful treatments might wrestle with key issues such as how to integrate multiple questions and outcomes and when in a sequence of studies (if a sequence is needed) key issues might be addressed.

EVALUATION OF THE STAGE MODEL

I have begun with the broader scientific agenda because I believe that the best (but not the only) way to identify treatments that are effective in clinical work is to elaborate why therapy leads to change, how the changes come about, and what processes must be activated within treatment and the client to achieve change.[2] A central goal of the stage model is to bring therapy from clinical development to effective clinical application. The process is accomplished through pilot work, randomized controlled clinical trials, and tests of treatment in service delivery settings. I believe this sequence will not achieve the goal the developing treatments that work well in clinical settings. That said, it is important to examine the model on its own merit by looking at what a particular stage is designed to accomplish, for instance. The emphasis of this article is on stage I and hence so is the emphasis of these comments.

Demands of Pilot Work

A difficulty with the model is the level of research sophistication that the pilot work (stage I) requires. Stage I can include attending to effect size, power, and sample size; developing a treatment manual, selecting a homogeneous sample; consulting with focus groups to examine the feasibility and acceptability of treatment, specifying client inclusion and exclusion criteria, and assessing processes of treatment. The differences between stage I and stage II can be large and convey clearly that the latter builds on the former in terms of rigor and demands on the investigator. Even so, stage I will require research-oriented and trained investigators, many of whom are not working in clinical settings and who are not exposed to seriously disturbed patients and who do not experience the impasses in treatment that often foster innovation.

Stage I might be a place to bring in clinicians to foster or formalize collaborative arrangements with researchers. More and more individuals in clinical practice are collecting usable data and information about their cases and generating or testing hypotheses about treatment (e.g., Clement, 1999; Maletzky, 1991; Marquis, 1991). A model designed to develop treatments might specify Stage I in ways that draw systematically on clinical work. It is before stage I that critically important ideas are likely to emerge. The preexperimental demonstrations and the pre-pilot study can reflect excellent bases for moving to a formal study. A model of treatment research that progresses from innovation to effective interventions in clinical work might integrate these prior steps.

Design Alternatives

The stage model focuses on group designs, statistical evaluation, and all that these entail (e.g., comparison groups, effect size, and control of extraneous variability so an effect can be demonstrated.) Group research designs are wonderful, but at the same time, stage I is an ideal place for single-case experimental designs (see Barlow & Hersen, 1984; Kazdin, 1982; Krishef, 1991). From a clinical perspective, single-case designs permit small-scale applications of treatment and allow the clinical investigator to tinker with treatment to develop effective variations. From a scientific perspective, these designs are true experiments and permit causal relations to be drawn between treatment and therapeutic change.

To highlight only one of the design options, in a multiple-baseline design, data are collected on two or more problem behaviors (e.g., specific symptom domains or a given problem across two or more situations or settings such as at home, at work, and in the community). Each behavior (or focus) reflects a separate baseline measure that is assessed on multiple occasions over time. After preintervention data are collected on each of the baselines, the intervention is then applied to alter one of the baselines while others continue to be observed. These other

baselines are roughly analogous to wait-list treatment control conditions, although the logic of the designs is somewhat different from treatment–control comparisons in the more familiar group research (Kazdin, 1998). The effect of an intervention is demonstrated experimentally by showing that each baseline changes as the intervention is introduced and as applied in a sequential fashion. From the standpoint of clinical work, the design is enormously helpful because one can apply the intervention on a small scale (one person, one behavior, behavior in one situation). Mediocre treatment effects can be redressed as needed in response to the data. Once an effective intervention is devised, it can be extended to other behaviors (or baselines) for that client. When one is trying to develop treatment and to devise a version that optimizes change, the ongoing data provide important feedback that the usual pre–posttreatment assessment of group designs does not afford.

Those unfamiliar with these designs and other single-case designs would reasonably state that the designs are not feasible in clinical work, raise problems of generalizing across clients, and are only useful with behavioral interventions. These concerns are not difficult to refute (Barlow, Hayes, & Nelson, 1984; Kazdin, 1998). More persuasively, the designs have been used clinically in scores of intervention studies, spanning multiple clinical problems, treatments, settings, and clients (e.g., infants to the elderly) (see Kazdin, 2001).

The stage model advocates true experiments, even in the pilot work. One cannot argue against the virtue of random assignment and other central features of such work. However, quasi-experiments often permit causal inferences about treatment effects. There are many examples of both single-case and group quasi-experiments in which the strength of the inferences is strong (Kazdin, 1998).

The stage model might make a bolder move in clarifying a broader portfolio of methodologies to achieve the goal of developing effective treatments. There is no reason to restrict pilot work or, indeed, treatment evaluation to group designs. Is it the case that single-case designs cannot generalize across individuals? Alas, the problems are elsewhere. The results of group designs do not necessarily generalize better. Indeed, statistical evaluation of group differences (the usual comparisons of means and estimation of effect sizes) does not reflect how many patients really improved or to whom the results might be generalizable. Indeed, the fixed treatment regimen of group designs and evaluation of progress after the fact (post and follow-up data), rather than ongoing assessment during treatment, are lamentable characteristics of group designs as they are currently used. These characteristics might limit greatly the impact of treatment for a given individual and for the broader population to whom one wishes to generalize. Is there no place for single-case designs and for quasi-experimental designs at the different stages in the stage model? There ought to be, and spelling these out could greatly advance intervention research and training of researchers.

Sequencing of Studies

The stage model poses a sequence of studies from pilot work, to RCTs, to tests in clinical settings. There is an obvious progression, and this sequence is logical. At the same time, the sequence raises questions. First, addressing a broader scientific agenda of therapy research may not be easily achieved as a linear progression through stages. There is a critically important iterative process that is central to scientific research. For a stage model, it would be useful to see a flow chart that not only moves forward in a linear way, but conveys some of the tributaries that are important or likely to emerge and how they feed back into the overall progression.

It is not merely the case that one must go back to the drawing board when treatment is not working. The real challenge is what to do next when treatment is working. There are many paths to take. For example, how and when does one move from an RCT back to a pilot study, or to a study of parametric variations of treatment, or to a study that dismantles treatment? Guidance and prioritization of these would be helpful. A test in clinical settings, I believe, is one of many options and not necessarily the most important one for achieving clinically effective treatments.

The challenge for a model of treatment research is delineating the process leading to knowledge because findings at one stage do not simply lead to progression to the next stage. For example, suppose cognitive therapy works very well in an RCT, but evidence in the study suggests that changes in cognitions were not involved in therapeutic change. Do we jump to a treatment study in a clinical setting? Do we work on the measures of cognitive processes to ensure that these processes are assessed in a way that is valid? Do we move back to the laboratory to

test multiple mechanisms that might be involved to understand what transpires in treatment? Probably we do them all. A model ought to address these tributaries to convey the nonlinear paths leading to the knowledge we need and perhaps the criteria for noting when a particular gap can be regarded as filled and when one ought to move on to other questions.

Clearly, there is a sequence of studies that can be readily identified. Thus, studies of a treatment package are logically prior to studies that attempt to dismantle treatment. The reason I listed previously many of the questions that guide therapy research was to convey the range of studies that are needed to understand the effects of treatment. The overarching question is, what do we need to know about therapy or a given treatment? The answer involves specifying the types of studies we need and the ways in which they can be conducted to obtain the answers. We need to end up in a particular place (e.g., knowledge about effects and the way in which they are achieved). The order in which studies are completed is less important than the accretion of a small number of studies on each of the critical questions.

CONCLUSIONS

A central goal of the stage model is to develop treatment so that it can be applied effectively in clinical work. The stage model is designed to move a treatment from pilot development to a trial in clinical settings. This is a critically important goal and has been discussed often with few constructive efforts to plan how it might be accomplished. The model is timely as well; there is continuing concern that moving from evidence-based treatment to evidenced-based clinical practice could well be eons away. The model addresses a central question and structures discussions of efficacy/effectiveness with a research plan.

My comments begin with a similar goal—namely, to develop treatments that can be and are used effectively in clinical work—but they underscore a slightly different path. If one is concerned with application (clinical extension), the scientific questions about the underpinnings of treatment are not esoteric, but just the opposite; the scientific questions about key facets of treatment, why and how treatments operate, become pivotal. To that end, any model or plan for research ought to be rather explicit on the steps needed to obtain the requisite knowledge. I would not expect the steps from pilot work, to RCTs, to an extension in practice to reach the goal effectively. This does not mean that one ought to wait to apply or test treatment clinically until one has all the answers; rather, one should develop a model that articulates the steps to obtain the requisite knowledge.

Tests that move from the laboratory to clinical settings raise their own problems. Effects of treatment in clinical practice are likely to be attenuated from those evident in clinical research. In the clinical situation there is variation in all sorts of factors including poor attendance by patients, comorbidity, inadequate resources, and disinterest in monitoring treatment fidelity, among a host of others. The final common pathways of these characteristics is to dilute treatment effects (reduce mean differences between treatment and control conditions) and to add variability (increase within-group standard deviations). These of course translate directly to reductions in obtained effect sizes and the likelihood of statistical differences.

There are different ways to redress the problem so as to permit better evaluations of clinical work and make extensions of well-studied treatments to clinical practice more feasible. One is to introduce some user-friendly characteristics of evaluation into clinical settings (e.g., Clement, 1999; Kazdin, 1993). Another way is to understand treatments much better than we do (i.e., understand why and how they work). All of the problems of administering treatments effectively in clinical settings are not erased by understanding treatment better. Yet one ought to be armed with knowing how treatment produces change, the critical components of treatment, and how to maximize their impact. No doubt there will be demonstrations of treatment in clinical settings that are shown to produce the desired effects. Will these be replicable? Perhaps some; the better guarantee to replication across settings and over time is understanding why treatment works and how to invoke the critical change processes. Without this work, tests of generality of treatment (stage III) are likely to have little generality.

The notion of a stage model is intriguing. If we want to go from intervention to application, what *are* the stages? Clearly a topic to address, well beyond the goals of the stage model as proposed and these comments, is dissemination and implementation into clinical practice. Suppose one of the currently empirically supported treatments (i.e., treatments shown to work in replications of laboratory-based studies) were shown to work in clinical settings. How then do we foster clinical use and ensure that the applications are carried out in the manner

intended or needed to effect change? There are few mechanisms for changing clinical practice and ensuring quality control of treatments that are administered. A broader plan is needed to move treatment from clinical innovation, through controlled trials, to adoption and dissemination. This article highlighted the priority of the scientific questions to ensure that a treatment is worth disseminating.

NOTES

1. The model is proposed in relation to behavioral therapies. The present comments are directed to psychotherapies in general (i.e., psychosocial interventions that are devised and implemented in the context of treatment).

2. There are other steps as well, such as connecting theory of clinical dysfunction to the foci of treatment and selecting moderating variables, as elaborated elsewhere (see Kazdin, 2000).

ACKNOWLEDGMENT

Completion of this article was facilitated by support from the Leon Lowenstein Foundation, the William T. Grant Foundation (98–1872–98), and the National Institute of Mental Health (MH59029).

REFERENCES

Barlow, D. H., Hayes, S. C., & Nelson, R. O. (1984). *The scientist-professional: Research and accountability in clinical and research settings*. New York: Pergamon.

Barlow, D. H., & Hersen, M. (1984). *Single-case experimental designs: Strategies for studying behavior change* (2nd ed.). New York: Pergamon.

Clement, P. W. (1999). *Outcomes and incomes: How to evaluate, improve, and market your practice by measuring outcomes in psychotherapy.* New York: Guilford.

Hoagwood, K., Hibbs, E., Brent, & Jensen, P. J. (1995). Efficacy and effectiveness in studies of child and adolescent psychotherapy. *Journal of Consulting and Clinical Psychology, 63,* 683–687.

Kazdin, A. E. (1982). *Single-case research designs: Methods for clinical and applied settings*. New York: Oxford University Press.

Kazdin, A. E. (1993). Evaluation in clinical practice: Clinically sensitive and systematic methods of treatment delivery. *Behavior Therapy, 24,* 11–45.

Kazdin, A. E. (1998). *Research design in clinical psychology* (3rd ed.). Needham Heights, MA: Allyn & Bacon.

Kazdin, A. E. (2000). *Psychotherapy for children and adolescents: Directions for research and practice*. New York: Oxford University Press.

Kazdin, A. E. (2001). *Behavior modification in applied settings* (6th ed.). Belmont, CA: Wadsworth.

Kendall, P. C., & Chambless, D. L. (Eds.). (1998). Special section: Empirically supported psychological therapies. *Journal of Consulting and Clinical Psychology, 66,* 3–167.

Kolvin, I., Garside, R. F., Nicol, A. E., MacMillan, A., Wolstenholme, F., & Leitch, I. M. (1981). *Help starts here: The maladjusted child in the ordinary school.* London: Tavistock.

Krishef, C. H. (1991). *Fundamental approaches to single-subject design and analysis*. Malabar, FL: Kreiger.

Lonigan, C. J., & Elbert, J. C. (Eds.). (1998). Special issue on empirically supported psychosocial interventions for children. *Journal of Clinical Child Psychology, 27,* 138–226.

Maletzky, B. M. (1991). *Treating the sexual offender.* Newbury Park, CA: Sage.

Marquis, J. N. (1991). A report on seventy-eight cases treated by eye movement desensitization. *Journal of Behavior Therapy and Experimental Psychiatry, 22,* 187–192.

Meyers, A. W., Graves, T. J., Whelan, J. P., & Barclay, D. (1996). An evaluation of television-delivered behavioral weight loss program. Are the ratings acceptable? *Journal of Consulting and Clinical Psychology, 64,* 172–178.

Nathan, P. E., &. Gorman, J. M. (Eds.). (1998). *Treatments that work*. New York: Oxford University Press.

Newman, M. G., Consoli, A., & Taylor, C. B. (1997). Computers in assessment and cognitive behavioral treatment of clinical disorders: Anxiety as a case in point. *Behavior Therapy, 28,* 211–235.

Onken, L. S., Blaine, J. D., & Battjes, R. (1997). Behavioral therapy research: A conceptualization of a process. In S. W. Henggeler & R. Amentos (Eds.), *Innovative approaches from difficult to treat populations* (pp. 477–485). Washington, DC: American Psychiatric Press.

Snyder, C. R., & Ingram, R. E. (Eds.). (2000). *Handbook of psychological change: Psychotherapy processes and practices for the 21st century.* New York: John Wiley & Sons.

Stone, T., & Darlington, G. (2000). *Pills, potions, and poisons: How drugs work*. New York: Oxford University Press.

Szapocznik, J., Rio, A., Murray, E., Cohen, R., Scopetta, M., Rivas-Vasquez, A., Hervis, O., Posada, V., & Kurtines, W. (1989). Structural family versus psychodynamic child therapy for problematic Hispanic boys. *Journal of Consulting and Clinical Psychology, 57,* 571–578.

Wolfe, D. A. (1999). *Child abuse* (2nd ed.). Newbury Park, CA: Sage.

Received July 14, 2000; accepted July 21, 2000.