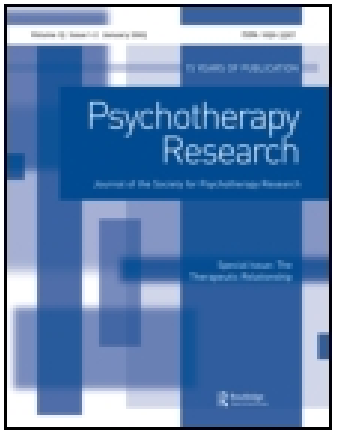

Psychotherapy Research
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/tpsr20

Understanding how and why psychotherapy leads to change

Alan E. Kazdin [a]
[a] Department of Psychology , Yale University , New Haven, CT, USA
Published online: 22 Sep 2009.

Psychotherapy Research, July–September 2009; 19(4–5): 418–428

PSYCHOTHERAPY RESEARCH METHODS

Understanding how and why psychotherapy leads to change

ALAN E. KAZDIN

Department of Psychology, Yale University, New Haven, CT, USA.

(Received 6 June 2008; revised 27 August 2008; accepted 28 August 2008)

Abstract
The rich research literature on psychotherapies for children, adolescents, and adults has identified several interventions that can ameliorate or effect significant change in psychiatric disorders and a variety of social, behavioral, and emotional problems and can enhance medical outcomes and recovery. After decades of psychotherapy research and thousands of studies, there is no evidence-based explanation of how or why even the most well-studied interventions produce change, that is, the mechanisms through which treatments operate. This article discusses central requirements for demonstrating mediators and mechanisms of change. Also presented are promising lines of work to identify mediators and mechanisms, ways of bringing to bear multiple types of evidence, recommendations to make progress in understanding how therapy works, and conceptual and research challenges in evaluating mediators and mechanisms.

Keywords: mediators and mechanisms of psychotherapy

Meta-analyses and narrative reviews of well-controlled studies have indicated that many forms of psychotherapy for children, adolescents, and adults lead to therapeutic change (e.g., Kazdin & Weisz, 2003; Lambert, 2004; Nathan & Gorman, 2007). Multiple questions remain, including the extent and indeed whether many treatments make a palpable difference in the lives of those treated, whether some treatments are more effective than others, and how we can harness the many factors that influence outcome. Arguably the most pressing question is how therapy leads to change. Currently, we do not know the reasons, although many ideas have been proposed. The focus of this article is on mediators and mechanisms of therapeutic change. Central to this article is the thesis that, with isolated exceptions, we do not know why or how therapies achieve therapeutic change, the requisite research to answer the question is rarely done, and fresh approaches are needed in conceptualization and research design.

Understanding why and how therapy leads to change are important for several reasons. First, there are scores of therapies with evidence that they produce change. Understanding mechanisms could bring order and parsimony if a few key mechanisms were identified to explain many treatments. Second, by understanding the processes that explain therapeutic change, one ought to be better able to optimize change. If we know how changes come about, perhaps we can identify better, different, or more strategies that trigger critical change processes. Third, to optimize the generality of treatment effects from research to practice, we want to know what is needed to make treatment work and what components must not be diluted to achieve change. Arming practitioners with evidence-based treatments is a valuable advance, but it would be even better if we could convey what facets are critical to include. Fourth, understanding how therapy works can help identify moderators of treatment (i.e., variables on which the effectiveness of a given treatment may depend). Understanding the processes through which treatment operates can help sort through those facets that might be particularly influential in treatment outcome and permit better selection and triage of suitable patients.

Although we do not know why therapy leads to change, this is not because of a lack of attention or

Correspondence concerning this article should be addressed to Alan E. Kazdin, Department of Psychology, Yale University, P.O. Box 208205, New Haven, CT, USA 06520-8205. E-mail: Alan.Kazdin@Yale.Edu

ISSN 1050-3307 print/ISSN 1468-4381 online
DOI: 10.1080/10503300802448899

hypotheses. Decades of research on therapy processes have identified many features of the client, the therapist, and their interaction as well as treatment activities that predict therapeutic outcome (see Orlinsky, Rønnestad, & Willutzki, 2004). Overlapping with and drawing from that research has been extensive discussion of constructs that might explain treatment effects, including the therapeutic relationship, catharsis, therapist warmth, learning, change in expectations, mastery, common factors among different therapies, and others (e.g., Lambert & Ogles, 2004; Wampold, 2001). Despite the attention and rather vast literature, there is little empirical research to provide an evidence-based explanation of precisely why treatment works and how the changes come about. In using the term *evidence-based explanation,* I am referring to replicated findings that convey the mechanisms responsible for change for a given treatment and how these mechanisms operate to produce symptom improvement.

Conceptual and Definitional Issues

Several interrelated and overlapping concepts are important to distinguish (Table I). It is useful to begin with cause or causal relation. A randomized controlled trial may show that treatment compared with no treatment leads to therapeutic change. From the demonstration we can say that the treatment *caused* the change, as that term is used in science. Demonstrating a cause does not say why the intervention led to change or how the change came about. To evaluate how change comes about, research often looks at mediators.

Table I. Key Terms and Concepts

Cause: A variable or intervention that leads to and is responsible for the outcome or change.

Mediator: An intervening variable that may account (statistically) for the relationship between the independent and dependent variables. Something that mediates change may not necessarily explain the processes of how change came about. Also, the mediator could be a proxy for one or more other variables or be a general construct that is not necessarily intended to explain the mechanisms of change. A mediator may be a guide that points to possible mechanisms but is not necessarily a mechanism.

Mechanism: The basis for the effect (i.e., the processes or events that are responsible for the change; the reasons why change occurred or how change came about).

Moderator: A characteristic that influences the direction or magnitude of the relationship between an independent and a dependent variable. If the relationship between variables *x* and *y* is different for males and females, sex is a moderator of the relation. Moderators are related to mediators and mechanisms because they suggest that different processes might be involved (e.g., for males or females).

Note. Several sources can be consulted for further discussion of these concepts (e.g., Baron & Kenny, 1986; Campbell & Stanley, 1963; Kraemer et al., 1997; Kraemer, Stice, Kazdin, Offord, & Kupfer, 2001; Kraemer, Wilson, Fairburn, & Agras, 2002).

Mediator is a construct that shows important statistical relations between an intervention and outcome. This is an intervening construct that suggests critical processes about why change occurs. Even so, a mediator may not explain the precise process through which change occurs. The mediator might serve as a proxy for one or more variables with which it is correlated. More critical, the mediator may not and usually is not intended to explain precisely how the change comes about.

Mechanism refers to a greater level of specificity than mediator and reflects the steps or processes through which therapy (or some independent variable) actually unfolds and produces the change. Mechanism explains how the intervention translates into events that lead to the outcome. This is easily confused with the notion of mediation. For example, cognitions may be shown to mediate change in therapy. However, this does not explain specifically how the change came about (i.e., what are the intervening steps between cognitive change and reduced stress or anxiety). The goal is to understand the mechanisms of change; the study of mediators is often a first step.

Moderator refers to some characteristic that influences the direction or magnitude of the relation between the intervention and outcome. If treatment outcome varies as a function of characteristics of the patient or therapist (e.g., sex, ethnicity, temperament) or treatment delivery (e.g., individual vs. group treatment), these characteristics are moderators. I discuss these constructs and their relations later but first begin with some clarity about the focus.

Demonstrating Mediators and Mechanisms of Change

Considerable research has focused on mediators as intervening processes between the intervention and change. Mediators can be an important step along the path to identifying mechanisms, and both constructs are considered here. Drawing inferences about a mediator requires convergence of multiple criteria that act in concert. These are often discussed in statistical terms, but the conceptual underpinnings are no less critical.[1] Table II illustrates several criteria to establish mediation and then mechanism of action. The case for a mediator is built by a sequence of studies that address the criteria. Consider briefly two areas of psychotherapy research where mediators and the mechanisms of change are often discussed.

Table II. Requirements for Demonstrating Mediators and Mechanisms of Change

Strong association: Demonstration of a strong association between the psychotherapeutic (*A*) intervention and the hypothesized mediator of change (*B*) and an association between the proposed mediator (*B*) and therapeutic change (*C*). Strong might be measured in effect size or percentage of variance but usually is addressed statistically through mediation analyses that show how the relation between *A* and *C* depend on *B*.

Specificity: Demonstration of the specificity of the association among the intervention, proposed mediator, and outcome. Ideally, many plausible constructs do not account for therapeutic change, with the exception of one, which strengthens the argument that the proposed construct mediates change.

Consistency: Replication of an observed result across studies, samples, and conditions (i.e., consistency in the relation) contributes to inferences about mediators. Inconsistency might result from operation of a moderator and not controvert interpretation of the critical construct. Yet consistency across studies greatly facilitates drawing inferences about whether a particular mediator may be involved.

Experimental manipulation: Direct experimental manipulation of the proposed mediator to show the impact on outcome (*C*).

Time line: Demonstrating a time line or ordering of the proposed mediator and outcome (i.e., the mediator changes before the outcome).

Gradient: Showing a gradient in which stronger doses or greater activation of the proposed mediator is associated with greater change in the outcome can help make the case for a particular mediator. No dose–response relation (e.g., a qualitative or on–off gradient) or a relation that is not linear does not refute the role of the construct but may make inferences more difficult to draw.

Plausibility or coherence: A plausible, coherent, and reasonable process that explains precisely what the construct does and how it works to lead to the outcome. The steps along the way (from construct to change) can be tested directly.

Note. See Kazdin (2007) for further details.

Therapeutic Alliance and Treatment Outcome

The therapeutic alliance (i.e., collaborative nature of the patient–therapist interaction, their agreement on goals, and the personal bond that emerges in treatment) is often posed as the explanation for why therapy works (i.e., a mediator and mechanism). This is understandable given the consistent finding that the stronger the alliance, the greater the therapeutic change (Horvath & Bedi, 2002; Orlinsky et al., 2004).

Studies that evaluate alliance during (e.g., early, middle) treatment often show that alliance predicts improvement in symptoms at the end of treatment. This is an important finding in its own right because prediction of change or lack of it might well guide the therapist in decision making about intervention strategies. However, showing that alliance predicts later symptom change by itself does not show that alliance plays a causal or mediational role in therapeutic change. Merely because symptoms are not assessed in the middle of treatment does not mean they have not already changed. Perhaps very early in treatment clients get a little better (some symptom improvement) and as a result form a positive alliance with the therapist. Symptom change and alliance need to be assessed early in treatment (and preferably at multiple points). Indeed, on more than one occasion, careful assessment shows that positive alliance may follow improvements in symptoms (Barber, Connolly, Crits-Christoph, Gladis, & Siqueland, 2000; DeRubeis & Feeley, 1990). More generally, the vast literature on alliance does not yet establish a causal relation between alliance and change; most studies do not meet the time line requirement.

Cognitions in Cognitive Therapy for Depression

There are very few forms of psychotherapy as well established as cognitive therapy (CT) for unipolar depression among adults (Hollon & Beck, 2004). But why does CT work, that is, through what mediators or mechanisms? CT is designed to change these cognitions and in the process change depression. The relation of cognitions and cognitive change in treatment to therapeutic change has been studied in different ways by assessing symptom change and cognitive change at the end of treatment and showing that one shares variance with the other or by evaluating whether cognitions assessed early or in the middle of treatment correlate with subsequent therapeutic change (e.g., DeRubeis et al., 1990; Kwon & Oei, 2003). In both of these methods, the time line problem is unresolved (i.e., we do not know the ordering of cognitive change and symptom change). This issue is similar to the concern raised in relation to alliance, namely that in the vast majority of studies symptom change may have preceded or occurred concurrently with cognitive changes.

Unlike the research on alliance, perhaps one can say a bit more about mediators of cognitive therapy. Tests of mediation and evaluation of therapeutic changes quite early in the course of treatment suggest that improvements can readily occur without changes in cognitions or in advance of implementing cognitive change strategies in treatment (e.g., Burns & Spangler, 2001; Tang & DeRubeis, 1999). Perhaps we can state more confidently now than before that whatever may be the basis of change with CT, changes in cognitions, as originally proposed, are not necessary conditions for therapeutic change.

General Comments

I have highlighted two examples to convey how a key criterion, establishing a time line, is not routinely met in otherwise well-studied areas where mediators and mechanisms are discussed. The examples were used to illustrate this single point rather than to review comprehensively the respective literatures.

Although I focused on the time line problem, other concerns in relation to these literatures could be illustrated by applying all of the criteria. For example, we do not have a clear picture or set of studies that test how the putative mechanisms of psychotherapy unfold in such a way as to alter symptoms. In relation to plausibility and coherence of the mechanism–outcome relation, precisely what happens that leads to symptom change? Through what process or sequence of events along any dimensions (cognitive processes, neurotransmitters, stress) does alliance lead to reductions in depression, anxiety, or feelings that life is meaningless? The time sequence problem is more basic, but how does one gets from "My therapist and I are bonding" to "My marriage, anxiety, and tics are better"? This is a leap with the intervening steps unspecified or untested, at least to my knowledge. The steps are not academic. If we could identify the steps, there may be other ways to activate them than through alliance alone. Also, we might identify novel moderators related to the mechanisms that help us select individuals likely to vary in responsiveness to the intervention.

What is needed here is testable theory about the intervening steps and how they unfold. Emphasis here is on "testable" rather than merely plausible global accounts. For example, it may be that the relationship changes patient affect of discussing stressful topics, and as these topics are discussed they become less arousing or stressful. Perhaps the effects extend beyond the treatment setting. I raise this sequence because it has been proposed decades ago with evaluation of therapist actions on client arousal within the session (e.g., Dittes, 1957a, 1957b). Extinction models have figured prominently in conceptualizations of treatment of anxiety and arousal and have the advantage of being able to draw on rich human and nonhuman animal laboratory research.

Paths to Identifying and Elaborating Mediators and Mechanisms

We begin with the questions, what is a possible mechanism and how might it operate? Several strategies and tests might be brought to bear to reach the answers. We can build the case by meeting the requirements outlined previously (see Table II).

Meticulous Description

In research one can readily distinguish description (what is happening) from explanation (why it is happening or through what forces, processes, or mechanisms). Depending on the detail, level of analysis, and sequence of moving from one to the other, description can blend into explanation. For example, cross-sectional and longitudinal studies and research with human and nonhuman animals have established a causal role between cigarette smoking and lung cancer. However, this leaves unexplained the mechanisms (i.e., the processes through which lung cancer comes about). The mechanism has been uncovered by describing what happens in a sequence from smoking to mutation of cells into cancer (Denissenko, Pao, Tang, & Pfeifer, 1996). A chemical (benzo[*a*]pyrene) found in cigarette smoke induces genetic mutation at specific regions of the gene's DNA that is identical to the damage evident in lung cancer cells. This finding is considered to convey precisely how cigarette smoking leads to cancer at the molecular level. This is an example of where "the what" (description) can be sufficiently fine grained as to convey "the how."

In therapy, proposed mechanisms might encompass such constructs as the therapeutic relationship. Research must then go beyond the demonstrated correlation and even the predictive portion (i.e., on the assumption that the time line can be firmly established). One way to move closer to understanding mechanisms would be to describe social interaction outside of the context of therapy in relation to neurological or other biological indices (e.g., Adolphs, 2003; Meyer-Lindenberg et al., 2005). What changes take place in social interaction? There is still a huge leap between these descriptions and explaining how a relationship in therapy leads to symptom change, but this is a start and moves beyond where we are today in the psychotherapy literature.

Given the examples, it is critical to underscore that meticulous description does not require evaluation of biology or biological correlates. For example, the emergence and escalation of aggressive child behavior in the home have been carefully described in several studies (e.g., Patterson, 1982; Patterson, Reid, & Dishion, 1992; Reid, Patterson, & Snyder, 2002). By directly observing family interaction in the home, coding multiple behaviors and interchanges, and evaluating conditional probabilities associated with particular interchanges, Patterson et al. have identified causal and reciprocal causal sequences of how parent–child interactions foster and escalate aggressive child behavior.

Moderators as a Path to Identifying Mediators and Mechanisms

Moderators can play a direct role in elaborating mediators and mechanisms of action. Consider an example of the effect of experience during childhood on subsequent criminal behavior, where a genetic characteristic is a moderator. As is well known, children with a history of physical abuse are at risk for later antisocial behavior (Child Welfare Information Gateway, 2006). Most people who are abused as children do not engage in antisocial behavior. A genetic characteristic moderates the relationship. Abused children with a genetic polymorphism (related to the metabolism of serotonin) have much higher rates of antisocial behaviors than those without this polymorphism (Caspi et al., 2002). Among boys with the allele and maltreatment, 85% developed some form of antisocial behavior (diagnosis of conduct disorder, personality assessment of aggression, symptoms of adult personality disorder, or court conviction of violent crime) by the age of 26. Individuals with the combined allele and maltreatment constituted only 12% of the sample but accounted for 44% of the cohort's violent convictions. Further research has replicated and extended the finding by noting that parent neglect as well as abuse in conjunction with the polymorphism increase risk for later conduct problems and violence (Foley et al., 2004; Jaffee et al., 2005).

So far, this is a fascinating illustration of moderation. However, closer scrutiny hints at mechanism. Caspi et al. (2002) looked at the allele for monoamine oxidase A (MAO-A) because

> The gene that encodes the MAO-A enzyme, which metabolizes neurotransmitters, is linked with maltreatment victimization and aggressive behavior.
> A rare mutation causing a null allele at the MAO-A locus in human males is associated with increased aggression.
> Animal gene knockout studies show that deleting this gene increases aggression.
> Restoring this gene expression decreases aggression.

In one sense we have identified a moderator: The influence of an independent variable (abuse in the home) and outcome (antisocial behavior years later) is moderated by some other characteristic or variable (MAO-A allele). Clearly, we have much more because the work and what it generated are beginning to point to possible genetic and molecular underpinnings. We do not know how the allele and abuse traverse specific steps from *A* to *Z* in which aggression emerges, but we are getting closer. For example, findings show the neural mechanisms through which the genetic influence is likely to operate (Meyer-Lindenberg et al., 2006). The MAO-A allele is associated with diminished brain circuitry related to impulse control that would promote aggression.

The type of moderator work illustrated here departs from the usual moderator research in psychotherapy. In the illustration, the moderator was identified based on considering mechanisms that might be involved. Theory about potential mechanisms, prior correlational evidence (abuse and victimization), and nonhuman animal studies served as background. In much of treatment research, and moderator research in clinical psychology more generally, moderators of convenience are used in which information routinely gathered (e.g., socioeconomic standing, age, ethnicity) is evaluated. There is little sound theory behind the research or predictions that derive from proposing precisely what facets of the moderator might be important in explaining the relation. Thus, there is a vast literature with analyses showing that boys and girls, younger versus older, and this ethnic group versus that ethnic group differ. This is fine as a start, but it needs to be pursued to convey why the moderator makes a difference. Moderation can lead to insights about mediation as the example of aggression shows, but it requires tests of ideas about what the mechanisms are or could be.

Direct Manipulation

Direct manipulation of a proposed mechanism is a powerful way to move our understanding forward. Consider the work on fear conditioning and psychotherapy. There have been decades of research on Pavlovian conditioning of fear in human and nonhuman animals. Conditioning as an explanation of fear acquisition and extinction as an explanation of fear reduction or elimination are useful paradigms for the processes that might be involved in treatment. Conditioning and extinction of fear depend on a particular receptor in the amygdala (*N*-methyl-d-aspartate; see Davis, Myers, Chhatwal, & Ressler, 2006). In nonhuman animal research, chemically blocking the receptor shortly before extinction training blocks extinction, demonstrating a dose–response relation. Blocking the receptor after extinction training also blocks extinction, which suggests that the consolidation process can be interrupted. A compound (D-cycloserine) binds to the receptor and makes the receptor work better (i.e., enhances extinction when given before or soon after extinction training).

The laboratory research has moved to therapy trials where exposure therapy, based on an extinction model, was evaluated to test whether enhancing a mechanism of extinction would improve treatment outcome. Activation of the critical receptor (with D-cycloserine vs. a placebo) has improved the therapeutic effects for individuals with acrophobia (fear of heights) as reflected on paper-and-pencil, physiological, and behavioral measures (Ressler et al., 2004). The enhanced outcomes have been replicated for social anxiety and obsessive-compulsive disorder (e.g. Hofmann et al., 2006; Wilhelm et al., 2008).

Needless to say, more work is needed. Experimentally induced fear in nonhuman animals in the laboratory varies in multiple ways from human fears with diverse and incompletely elaborated causes and moderators. That said, there are few examples like this in which identifying mechanisms of change in animal research has been explicitly extended to clients in the context of therapy. Understanding mechanisms of learning and extinction, but also memory, belief, persuasion, control, stress alleviation, anticipation, and so on, are within empirical reach in a similar way. Once such mechanisms are studied, potential targets can be identified, with a similar paradigm of manipulating the mechanisms.

Converging Lines of Work

Multiple lines of evidence are likely to be needed to converge on precisely what the mechanism is. Some of the examples I have provided focus on moderators and mechanisms and underpinnings that are biological. This is not a coincidence: The technological advances for studying biological processes are remarkable, and in some cases processes (e.g., neurotransmitter or synapse activity) can be observed in real time. Studying mediators and mechanisms and key theses of this article have nothing inherently to do with biology. The focus on mechanisms and the convergence of multiple lines of work can be gleaned from studying psychological processes, as mentioned previously in the context of family interaction and child aggression.

An active area of research focuses on the interplay of biological and psychological processes such as the neurological correlates and underpinnings of self-control, fear and flight, decision making, and learning (e.g., Baumeister, 2008; Marchiori & Warglien, 2008; Sanfey, 2007). This is not the reductionism of yesteryear, in which social scientists feared that a biological interpretation or account was an effort to replace or make unnecessary any psychological account. Indeed, basic biological research (e.g., animal models of addiction) suggests antecedents (e.g., impulsivity as a vulnerability to addiction) that might well prove to be targets for psychotherapeutic intervention (see Belin, Mar, Dalley, Robbins, & Everitt, 2008). The interplay of psychological and biological processes, the importance of their interdependence, and using each to influence the other hold great potential in enhancing our understanding of therapy and therapeutic activities that alter individual functioning.

Recommendations for Research

Use Theory as a Guide

Investigation of mediators and mechanisms of therapy can be improved in several ways. The guiding question for treatment research is, how does treatment achieve change? The answer may involve basic psychological processes (e.g., memory, learning, information processing) or a broader theory (e.g., motivation). What is needed further is greater specificity in conceptualizing not only the critical construct but also how that operates to produce symptom change. We need more than tests of mediation to understand mechanisms. Mediation tests of plausible constructs can provide a screening device of sorts to identify potential avenues to be pursued in a fine-grained fashion.

It would be helpful for intervention research to identify candidate mediators and mechanisms or plausible constructs that would explain or account for (statistically) therapeutic change, manipulate the proposed mechanism, assess to ensure it has been manipulated, and then evaluate change. For example, in relation to tobacco use among teenagers, several mediators that may serve as useful targets have been identified, including coping skills of the youth, peer influences, and availability of tobacco, among others (MacKinnon, Lockwood, Hoffman, West, & Sheets, 2002). The targets can be the focus of intervention. If one of these targets leads to change in tobacco use, this would serve as an excellent basis for further work to understand exactly how the influence produces change.

Include Measures of Potential Mediators in Treatment Studies

The mediator or mechanism ought to be specified so it can be measured. Studies occasionally include such measures of mediators (Hofmann, 2000; Weersing & Weisz, 2002), although most of the time their administration has not allowed evaluation of time lines. Yet measures are available. More fine-grained analyses will be needed to study the unfolding of processes over time and how change in some process results in symptom change. As a prerequisite to

understanding, assessments of potential mediators ought to be included in treatment studies.

Assess More than One Mediator or Mechanism

The accumulation of evidence would profit from the assessment of more than one mediator in a given study. It is rare that one mediator is studied, and hence there may be little value in raising the bar even higher by recommending the assessment of two or more mediators. Yet the assessment of multiple mediators in a given study has enormous benefits. If two or more mediators are studied, one can identify whether one is more plausible or makes a greater contribution to the outcome. In addition, the assessment of multiple potential mediators within individual studies is economically efficient, given the tremendous amount of time and resources needed for any treatment investigation. Across many studies, some mediators may repeatedly emerge as possible contenders while others fall by the wayside.

Establish the Time Line of the Proposed Mediator and Outcome

It is important to establish that the proposed mediator changes before the "outcome." The time line has two requirements: (a) The proposed mediator must be assessed before the proposed outcome and (b) the outcome must also be assessed early to ensure the mediator has, in fact, changed before the outcome. Even during the middle of treatment, long before the investigator may be interested in therapeutic change, it is quite possible that improvements occur before change in the putative mediator.

Assessment is the main change needed in research. Assessment on multiple occasions during treatment can provide information on the time line of mediators and outcomes and the possibility of bidirectional changes (i.e., each one influences the other in some way and at different points). Assessment on a session-by-session basis (i.e., every occasion over the course of treatment) permits evaluation of the mediator of change and symptom reduction and considers individual differences in the course of these changes.

Examine Consistencies Across Different Types of Studies

Understanding mediators and mechanisms through which therapeutic change occurs could profit from different types of studies, beyond those that might be construed as therapy research. Conclusions from these studies may be consistent and converge in making a particular process plausible. Nonhuman animal laboratory research might be relevant, as already illustrated in fear conditioning research. Therapeutically relevant phenomena (e.g., attachment, separation, social support) can be studied in animal research to identify processes (e.g., changes in the structure or function of the brain) and their consequences in behavior. These, in turn, might direct research to plausible underpinnings to support a conceptual view of the mechanism of therapeutic change. Such tests, far removed from therapy settings, provide important tests of principle. For example, maternal caregiving behaviors (e.g., nursing, licking, grooming) among rats influence the responsiveness to stress in the offspring; the effects can be seen in behavioral as well as neurological and endocrine responses (e.g., Champagne, Francis, Mar, & Meaney, 2003; Pruessner, Champagne, Meaney, & Dagher, 2004). This might well be pertinent to understanding stress, coping, and interventions designed to ameliorate stress. Are their therapeutic endeavors (e.g., coping skills training, positive regard of the therapist, planned activities outside of therapy) that can have parallel or similar neurological and endocrine impact as well as reduce stress?

Naturalistic studies might be very useful too. If one is proposing a mediator of change, is there a sample, population, or setting in which this mediator may be expected to vary naturally (i.e., without investigator intervention)? For example, if changing parenting style is proposed to explain why a parent- or family-based treatment of a child clinical problem is effective, naturalistic studies examining families with and without these practices and the short- and long-term child behaviors with which these are associated are relevant.

Among naturally occurring instances of the process or construct, can group differences be identified? For example, naturalistic studies of "normal" mothering have revealed that stress reactivity in human infants is influenced by maternal caregiving (e.g., sensitivity, availability, lack of intrusiveness) during routine activities (e.g., feeding, meal preparation), very much in keeping with the nonhuman animal research highlighted previously (Hane & Fox, 2006). Low-quality caregiving was associated with greater stress reactivity of their infants (e.g., fearfulness, more right frontal brain asymmetry), an effect that could not be explained by infant temperament. Caregiving in relation to stress response and reactivity behaves in a similar way across different research paradigms and draws attention to mediators or mechanisms that might be pertinent to therapy (e.g., trauma, stress, coping). Naturalistic studies by themselves may not permit strong causal conclusions. And we as researchers are all too quick to note that correlational studies are inherently limited. Yet

observing processes that may be operative in the natural environment and their short- and long-term correlates can be very useful for both generating and testing hypotheses about mediators and mechanisms. Also, the scientific case for mechanisms is derived from converging clues, and in that context tests from correlational research can provide unique contributions not otherwise available.

Laboratory studies of therapeutic processes often are viewed with ambivalence because they do not show whether treatment works in real-life settings. This concern has been voiced in the context of studies evaluating treatment outcome as their primary, if not exclusive, focus. Controlled studies of therapy in research rather than clinical settings are now more important than ever in relation to mechanisms of change. For example, studies of psychosocial interventions (e.g., persuasion, advice) in a therapeutic or quasi-therapeutic context (e.g., focusing on affect-laden topics or clinical topics) on brain functioning (e.g., via neuroimaging) would be very valuable to document the scope of impact and the systems involved. This is not therapy but pursuit of leads for mechanisms that might be involved with therapeutic interventions. The careful control afforded such research is precisely what is needed to identify mediators and mechanisms.

Intervene to Change the Proposed Mediator or Mechanism

An excellent strategy is to conduct an experiment in which the proposed mediator is, in fact, altered or varied across groups, as mentioned in the treatment studies cited previously on extinction of fear. Groups randomly composed might be assigned to low, high, and medium levels of a proposed mediator (as a general concept) or mechanism (as a more specific set of steps expected to lead directly to the outcome). Strong support would be evident from findings that outcome varies directly as a function of levels of the manipulated dose.

A variation of the intervention approach—*therapy knockout studies*—is worth distinguishing. The term draws from genetic work (e.g., gene knockout studies with mice) in which a particular gene is omitted or altered, and the effects are evident on behavior or some other facet suspected to be controlled by the gene. Similarly, if the psychotherapy investigator believes or theory predicts that a specific mechanism accounts for change, it would be useful to provide the therapy with an added intervention that is designed to "knock out" (inactivate) the mechanism. If role-play, practice, or warm, fuzzy relations are critical to the technique, give two variations of the treatment: the original and the original with an effort to inactivate the mechanism. In any single study, supportive evidence that treatment worked only when the mechanism was allowed to operate could be explained in multiple ways. Even so, this evidence would be a superb addition to accumulating evidence.[2]

Special challenges and obstacles

There are multiple challenges in considering mediators and mechanisms that extend beyond a few changes in designs or measurement strategies. First, my comments have implied a simple model in which a single mechanism leads to a single outcome or the effects are strong, simple, clear, and uniform. Yet a single influence can produce multiple outcomes. For example, cigarette smoking leads to several physical and psychological conditions. In some of these, we know there is a causal relation and the mechanism; in others we know of increased risk. The pervasiveness of the influence of smoking on so many conditions can introduce complexities in the search for mechanisms because many biological systems are involved. There may be multiple and different mechanisms for the single agent but different outcomes.

Second and related, similar outcomes may be reached through multiple paths. Thus, we do not expect to see all people with a particular characteristic to have achieved that through the same path. The paths may reflect similar mechanisms activated by different experiences or different mechanisms. For example, we have learned that schizophrenia and autism might be caused by many different gene variants, some of which are spontaneous (rather than inherited) and rare (Bakkaloglu et al., 2008; Walsh et al., 2008). As a more familiar example, low IQ could result from genetic, prenatal, cultural, and postnatal toxic (e.g., lead) influences. Thus, a single outcome has many paths. Essential to work on mediators and mechanisms is distinguishing different courses or paths and moderating influences. Looking for one explanation or mechanism for one group, one therapy, or one outcome may yield little.

Third, we often think in and solely test for linear relations, although many relations are nonlinear. For example, cholesterol and risk for heart disease are positive and linear; higher cholesterol increases risk. Cholesterol and stroke are U shaped (i.e., nonlinear), so that low and high cholesterol increase risk. Nonlinear relations propose a challenge in the sense that dose–response relation (as a linear function) is one clue on the path toward mechanisms, although it is not essential. Looking only for linear relations could mask reliable patterns of mediator-outcome

relations. One can test for nonlinear dose–response relations.

Fourth, even if all patients change on the basis of the identical mechanisms, the timing and patterns of change may vary (e.g., Stulz & Lutz, 2007). Some patients may make rapid or sudden gains at a particular point in treatment (e.g., Busch, Kanter, Landes, & Kohlenberg, 2006; Lutz & Tschitsaz, 2007; Tang & DeRubeis, 1999). One could say that at a given point some have and some have not made change in some qualitative or categorical fashion. Alternatively, one could consider that the point of therapeutic change for all individuals is normally distributed with a mean and standard deviation. In either scenario (sudden gains but not at the identical point or normally distributed changes across several points), assessment of the mechanism is a challenge. Assessment of the mechanism at any one or two points in a study may not capture when change in the mechanism has occurred for each individual. A challenge for research is ensuring that one can evaluate mechanism and change that may vary in course among individuals.

Finally, it is possible that the mediator or mechanism of change in psychotherapy varies as a function of a moderator variable. Searching for moderators (a priori or post hoc), testing them (statistical power from dividing of the sample into subgroups), and interpreting them (e.g., is the moderator a proxy for some other variable?) have their own special challenges. Rather than looking for main effects of an intervention and a uniform mechanism of change, we may need to identify and characterize subgroups, very much in the way that genetic researchers often profit from looking at special groups and individual outliers.

Conclusions

There has been enormous progress in psychotherapy research. This has culminated in recognition of several treatments that have strong evidence in their behalf. Despite this progress, research advances are sorely needed in studying the mediators and mechanisms of therapeutic change. It is remarkable that after decades of psychotherapy research we cannot provide an evidence-based explanation for how or why even our most well-studied interventions produce change. Extinction-based treatments, as highlighted, might be regarded as a possible exception, but this is hardly the focus in the vast majority of treatments evaluated in research and used in clinical practice.

Extensive attention is accorded statistical analyses and the conditions that need to be met in order to conclude that mediation has been demonstrated (e.g., Kraemer et al., 2008; MacKinnon, 2008). Statistical analyses, as I noted previously, are interconnected with the conceptual and design points of this article. I have omitted statistical analyses to cast a brighter light on what we ought to look for in the conceptualization and design of studies. The logic of what mediators and mechanisms are and the requirements for their elaboration are separate from data evaluation. One can still see examples of statistical tests of mediation applied in designs in which conclusions about mediators are not permitted because all variables (mediator, outcome) were assessed at the same time (cross-sectional study) or one variable (mediator) was assessed before the outcome, but we do not know whether the outcome, if assessed during the study, would have preceded, followed, or occurred simultaneously with change in the mediator.

The scientific study of mechanisms of change is certainly not an easy path on which to embark. A given treatment might work for multiple reasons. Just as there is no simple and single path to many diseases, disorders, and social, emotional, and behavioral problems (e.g., lung cancer, attention-deficit/hyperactivity disorder), there may be analogous complexity in mechanisms for a given treatment technique or therapeutic outcome. Two patients in the same treatment conceivably could respond for different reasons. The complexities are critically important to understand because the best patient care will come from ensuring that the optimal variation of treatment is provided. Understanding mechanisms of treatment is the path toward improved treatment.

Notes

[1] Statistical evaluation is intertwined with the conceptualization and evaluations of mediators (e.g., Baron & Kenny, 1986; Kenny, Kashy, & Bolger, 1998; Kraemer, Kiernan, Essex, & Kupfer, 2008; MacKinnon, 2008; Shrout & Bolger, 2002). The present article focuses on conceptual and methodological requirements. Not only have these received less attention, but applications of statistical methods occasionally have led to unwarranted conclusions in part because critical requirements (e.g., about the time line) have not been met. Also, statistical models can lead to quite different conclusions about mediation and hence deserve their own detailed assessment and presentation (Kraemer et al., 2008).

[2] This strategy bears similarity to but can be distinguished from dismantling therapy studies in which various components (procedures) of treatment are separated and provided to different groups to identify necessary, sufficient, and facilitative components of a treatment package. The focus is on procedures and what is needed for therapeutic change. Although this might be considered to imply mechanisms, it is not a direct test of mechanisms. The knockout strategy here focuses on critical mechanisms and manipulation of the mechanism.

References

Adolphs, R. (2003). Cognitive neuroscience of human social behaviour. *Nature Reviews Neuroscience*, *4*, 165–178.

Bakkaloglu, B., O'Roak, B. J., Louvi, A., Gupta, A. R., Abelson, J. F., Morgan, T. M., et al. (2008). Molecular cytogenetic analysis and resequencing of contactin associated protein-like 2 in autism spectrum disorders. *American Journal of Human Genetics*, *82*, 165–173.

Barber, J. P., Connolly, M. B., Crits-Christoph, P., Gladis, L., & Siqueland, L. (2000). Alliance predicts patients' outcome beyond in-treatment change in symptoms. *Journal of Consulting and Clinical Psychology*, *68*, 1027–1032.

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*, 1173–1182.

Baumeister, R. F. (2008). Free will in scientific psychology. *Current Perspectives in Psychological Science*, *3*, 14–19.

Belin, D., Mar, A. C., Dalley, J. W., Robbins, T. W., & Everitt, B. J. (2008). High impulsivity predicts the switch to compulsive cocaine-taking. *Science*, *320*, 1352–1355.

Burns, D. D., & Spangler, D. L. (2001). Do changes in dysfunctional attitudes mediate changes in depression and anxiety in cognitive behavioral therapy. *Behavior Therapy*, *32*, 337–369.

Busch, A. M., Kanter, J. W., Landes, S. J., & Kohlenberg, R. J. (2006). Sudden gains and outcome: A broader temporal analysis of cognitive therapy for depression. *Behavior Therapy*, *37*, 61–68.

Campbell, D. T., & Stanley, J. C. (1963). Experimental and quasi-experimental designs for research and teaching. In N. L. Gage (Ed.), *Handbook of research on teaching*. Chicago: Rand McNally.

Caspi, A., McClay, J., Moffitt, T. E., Mill, J., Martin, J., Craig, I., et al. (2002). Role of genotype in the cycle of violence in maltreated children. *Science*, *297*, 851–854.

Champagne, F. A., Francis, D. D., Mar, A.., & Meaney, M. (2003). Variations in maternal care in the rat as a mediating influence for the effects of development on behavior. *Physiology and Behavior*, *79*, 359–371.

Child Welfare Information Gateway. (2006). *Long-term consequences of child abuse and neglect.* Retrieved September 23, 2008, from www.childwelfare.gov/pubs/factsheets/long_term_consequences.cfm

Davis, M., Myers, K. M., Chhatwal, J., & Ressler, K. J. (2006). Pharmacological treatments that facilitate extinction of fear: Relevance to psychotherapy. *NeuroRx*, *3*, 82–96.

Denissenko, M. F., Pao, A., Tang, M., & Pfeifer, G. P. (1996). Preferential formation of benzo[a]pyrene adducts at lung cancer mutational hotspots in *P53*. *Science*, *274*, 430–432.

DeRubeis, R. J., Evans, M. D., Hollon, S. D., Garvey, M. J., Grove, W. M., & Tuason, V. B. (1990). How does cognitive therapy work? Cognitive change and symptom change in cognitive therapy and pharmacotherapy for depression. *Journal of Consulting and Clinical Psychology*, *58*, 862–869.

DeRubeis, R. J., & Feeley, M. (1990). Determinants of change in cognitive therapy for depression. *Cognitive Therapy and Research*, *14*, 469–482.

Dittes, J. E. (1957a). Extinction during psychotherapy or GSR accompanying "embarrassing" statements. *Journal of Abnormal and Social Psychology*, *54*, 187–191.

Dittes, J. E. (1957b). Galvanic skin response as a measure of patient's reaction to therapist's permissiveness. *Journal of Abnormal and Social Psychology*, *55*, 295–303.

Foley, D., Wormley, B., Silberg, J., Maes, H., Hewitt, J., Eaves, L., & Riley, B. (2004). Childhood adversity, monoamine oxidase A genotype, and risk for conduct disorder. *Archives of General Psychiatry*, *61*, 738–744.

Hane, A. A., & Fox, N. A. (2006). Ordinary variations in maternal caregiving influence human infants' stress reactivity. *Psychological Science*, *17*, 550–556.

Hofmann, S. G. (2000). Treatment of social phobia: Potential mediators and moderators. *Clinical Psychology: Science and Practice*, *7*, 3–16.

Hofmann, S. G., Meuret, A. E., Smits, J. A., Simon, N. M., Pollack, M. H., Eisenmenger, K., et al. (2006). Augmentation of exposure therapy with D-cycloserine for social anxiety disorder. *Archives of General Psychiatry*, *63*, 298–304.

Hollon, S. D., & Beck, A. T. (2004). Cognitive and cognitive behavioral therapies. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (5th ed) (pp. 447–492). New York: Wiley.

Horvath, A. O., & Bedi, R. P. (2002). The alliance. In J. C. Norcross (Ed.), *Psychotherapy relationships that work: Therapist contributions and responsiveness to patients* (pp. 37–69). New York: Oxford University Press.

Jaffee, S. R., Caspi, A., Moffitt, T. E., Dodge, K., Rutter, M., Taylor, A., & Tully, L. (2005). Nature × nurture: Genetic vulnerabilities interact with physical maltreatment to promote behavior problems. *Development and Psychopathology*, *17*, 67–84.

Kazdin, A. E. (2007). Mediators and mechanisms of change in psychotherapy research. *Annual Review of Clinical Psychology*, *3*, 1–27.

Kazdin, A. E., & Weisz, J. R. (Eds). (2003). *Evidence-based psychotherapies for children and adolescents*. New York: Guilford Press.

Kenny, D. A., Kashy, D. A., & Bolger, N. (1998). Data analysis in social psychology. In D. Gilbert, S. T. Fiske & G. Lindzey (Eds), *Handbook of social psychology* (Vol. 1 4th ed., pp. 233–265). New York: Oxford University Press.

Kraemer, H. C., Kazdin, A. E., Offord, D. R., Kessler, R. C., Jensen, P. S., & Kupfer, D. J. (1997). Coming to terms with the terms of risk. *Archives of General Psychiatry*, *54*, 337–343.

Kraemer, H. C., Kiernan, M., Essex, M., & Kupfer, D. J. (2008). How and why criteria defining moderators and mediators differ between the Baron & Kenny and MacArthur approaches. *Health Psychology*, *27*(Suppl.), S101–S108.

Kraemer, H. C., Stice, E., Kazdin, A. E., Offord, D. R, & Kupfer, D. J. (2001). How do risk factors work together? Mediators, moderators, independent, overlapping, and proxy-risk factors. *American Journal of Psychiatry*, *158*, 848–856.

Kraemer, H. C., Wilson, G. T., Fairburn, C. G., & Agras, W. S. (2002). Mediators and moderators of treatment effects in randomized clinical trials. *Archives of General Psychiatry*, *59*, 877–883.

Kwon, S., & Oei, T. P. S. (2003). Cognitive change processes in a group cognitive behavior therapy of depression. *Journal of Behavior Therapy and Experimental Psychiatry*, *34*, 73–85.

Lambert, M. J. (Ed.). (2004). *Bergin and Garfield's handbook of psychotherapy and behavior change* (5th ed.). New York: Wiley.

Lambert, M. J., & Ogles, B. M. (2004). The efficacy and effectiveness of psychotherapy. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (5th ed., pp. 139–193). New York: Wiley.

Lutz, W., & Tschitsaz, A. (2007). Sudden gains and losses in the treatment of patients with anxiety, depressive and comorbid disorders. *Zeitschrift fur Klinische Psychologie und Psychotherapie: Forschung und Praxis*, *367*, 298–308.

MacKinnon, D. P. (2008). *Introduction to statistical mediation analysis*. Mahwah, NJ. Erlbaum.

MacKinnon, D. P., Lockwood, C. M., Hoffman, J. M., West, S. G., & Sheets, V. (2002). A comparison of methods to test

mediation and other intervening variable effects *Psychological Methods*, *7*, 83–104.

Marchiori, D., & Warglien, M. (2008). Predicting human interactive learning by regret-driven neural networks. *Science*, *319*, 1111–1113.

Meyer-Lindenberg, A., Buckholtz, J. W., Kolachana, B., Hariri, A. R., Pezawas, L., Blasi, G., et al. (2006). Neural mechanisms of genetic risk for impulsivity and violence in humans. *Proceedings of the National Academy of Sciences*, *103*, 6269–6274.

Meyer-Lindenberg, A., Hariri, A. R., Munoz, K. E., Mervis, C. B., Mattay, V. S., Morris, C. A., & Berman, K. F. (2005). Neural correlates of genetically abnormal social cognition in William's syndrome. *Nature Neuroscience*, *8*, 991–993.

Nathan, P. E., & Gorman, J. M. (Eds). (2007). *Treatments that work* (3rd ed.). New York: Oxford University Press.

Orlinsky, D. E., Rønnestad, M. H., & Willutzki, U. (2004). Fifty years of psychotherapy process-outcome research: Continuity and change. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (5th ed., pp. 307–389). New York: Wiley.

Patterson, G. R. (1982). *Coercive family process*. Eugene, OR: Castalia.

Patterson, G. R., Reid, J. B., & Dishion, T. J. (1992). *Antisocial boys*. Eugene, OR: Castalia.

Pruessner, J. C., Champagne, F., Meaney, M. J., & Dagher, A. (2004). Dopamine release in response to a psychological stress in humans and its relationship to early maternal care: A positron emission tomography study using [C]raclopride. *Journal of Neuroscience*, *24*, 2825–2831.

Reid, J. B., Patterson, G. R., & Snyder, J. (Eds). (2002). *Antisocial behavior in children and adolescents: A developmental analysis and model for intervention*. Washington, DC: American Psychological Association.

Ressler, K. J., Rothbaum, B. O., Tannenbaum, L., Anderson, P., Graap, K., Zimand, E., et al. (2004). Cognitive enhancers as adjuncts to psychotherapy. Use of D-cycloserine in phobic individuals to facilitate extinction of fear. *Archives of General Psychiatry*, *61*, 1136–1144.

Sanfey, A. G. (2007). Social decision-making: Insights from game theory and neuroscience. *Science*, *318*, 598–602.

Shrout, P. E., & Bolger, N. (2002). Mediation in experimental and nonexperimental studies: New procedures and recommendations. *Psychological Methods*, *7*, 422–445.

Stulz, N., & Lutz, W. (2007). Multidimensional patterns of change in outpatient psychotherapy: The phase model revisited. *Journal of Clinical Psychology*, *63*, 817–833.

Tang, T. Z., & DeRubeis, R. J. (1999). Sudden gains and critical sessions in cognitive-behavioral therapy for depression. *Journal of Consulting and Clinical Psychology*, *67*, 894–904.

Walsh, T., McClellan, J. M., McCarthy, S. E., Addington, A. M., Pierce, S. B., Cooper, G. M., et al. (2008). Rare structural variants disrupt multiple genes in neurodevelopmental pathways in schizophrenia. *Science*, *320*, 539–543.

Wampold, B. E. (2001). *The great psychotherapy debate: Models, methods, and findings*. Mahwah, NJ: Erlbaum.

Weersing, V. R., & Weisz, J. R. (2002). Mechanisms of action in youth psychotherapy. *Journal of Child Psychology and Psychiatry*, *43*, 3–29.

Wilhelm, S., Buhlmann, U., Tolin, D. F., Meunier, S. A., Pearlson, G. D., Reese, H. E., Cannistraro, P., Jenike, M. A., & Rauch, S. L. (2008). Augmentation of behavior therapy with D-cycloserine for obsessive-compulsive disorder. *American Journal of Psychiatry*, *165*, 335–341.