Moderators and Mediators of Treatment Outcome for Youth With ADHD: Understanding for Whom and How Interventions Work

Stephen P. Hinshaw, PhD
Department of Psychology, University of California, Berkeley, Calif.

Objective To present data on moderators and mediators of treatment response from the Multimodal Treatment Study of Children With ADHD (MTA).

Methods Moderator variables (baseline factors that define subgroups with greater vs lesser intervention response) and mediator variables (factors occurring during treatment that explain how interventions ‘work’) are described with specific application to the outcomes of the MTA Study. Results Key moderator variables (comorbid anxiety disorder, public assistance, severity of attention-deficit/hyperactivity disorder, parental depressive symptomatology, IQ) and mediator processes (negative/ineffective parental discipline) are reviewed. Conclusions Treatment research in the future should explicitly consider the exploration of moderator and mediator variables, which can greatly aid the explanatory power of clinical trials and specify the critical next steps for intervention research.

Key words ADHD; mediator; moderator.

Although the chief aim of clinical trials regarding treatment strategies for any condition, including attention-deficit/hyperactivity disorder (ADHD), would appear to be a close examination of intervention effects on key outcomes, problems emerge when there is a sole focus on treatment-outcome linkages. First, it is difficult to know whether a single outcome measure exists, and the relative importance of ameliorating impairments versus reducing core symptomatology is often an open question. Second, a focus on immediate versus longer-term outcomes is a critical issue, particularly for a condition as persistent as ADHD. Third, no interventions for any mental health condition are universally effective, meaning that it is essential to gain understanding of which treatments work best for which particular kinds of individuals with a given disorder. Fourth, at the level of the independent variable (i.e., the treatment strategies under investigation), understanding the active ingredients that yield benefit is critical. Indeed, interventions for children and adolescents that appear to effect change because of certain program ingredients may actually produce benefits for very different reasons.

In short, outcome research will be enriched by investigation of relevant subgroups with particularly good (or poor) treatment response and by examination of relevant processes and mechanisms that yield clinically significant change. Without knowledge of the particular subgroups that respond best and worst to any particular treatment, and without systematic understanding of treatment mechanisms, research on effective treatments is bound to remain at a relatively primitive, descriptive level. Although research on adult conditions has a long history of the search for important treatment processes that may explain outcome, child and adolescent treatment research requires progress in this regard.

In this article, I focus on such issues in relation to ADHD, a prevalent childhood condition that yields substantial impairment and that clearly constitutes an issue of central importance for public health. I feature discussion of 2 essential constructs, termed moderator and mediator processes, which have been used and examined in basic psychological research, risk research related to psychopathology and—more recently—clinical trials studies. Moderator variables refer to characteristics of a treatment sample that may influence the outcomes of interest, thereby identifying subgroups with greater or lesser chances for positive response. Mediator variables signify processes occurring during treatment that explain...
how and why the intervention is exerting its effects. Certain mediators, in fact, may serve as underlying mechanisms of change, those processes that are causally responsible for bringing about improvements in outcome. Comprehension of these processes is essential to gain full benefit from the information gleaned from clinical trials. After describing these concepts in some detail and discussing their implications for treatment research, I cover what is known about key moderator variables and mediator processes that are operative in treatment research on youth with ADHD. In so doing, I provide methodologic and design recommendations for those interested in conducting future treatment studies in the field.

**Moderator and mediator variables**

Much of the research enterprise in the behavioral and biomedical sciences pertains to uncovering associations (correlations) between variables of interest. A correlation between 2 well-measured variables, however—for example, between a risk factor and a disease process—reflects the statistical association that exists across all cases in the sample, without respect to whether stronger or weaker associations pertain to certain subgroups. In addition, a simple correlation does not reveal the mechanisms and processes underlying the association. Indeed, the variables may be related not because of a causal relationship between them but because a third variable actually serves as an antecedent to both or because of an intricate chain of associations is at work.

In one of the most heavily cited social science articles of the past 2 decades, Baron and Kenny explained moderator and mediator variables with respect to research in social psychology, providing expert guidelines on how to test for their presence. Although full explanation of the complex nature of detecting and analyzing for moderator and mediator processes is beyond the scope of the present article, several issues are salient.

First, moderation implies that relationships of interest differ significantly in 2 (or more) subpopulations, which are defined by the moderator variable. Establishing the presence of a moderator effect therefore requires that there is a statistical interaction between the moderator variable and the predictor variable, such that at different levels of the moderator, the predictor-criterion relationship is different in magnitude—or even in sign. For example, it may be the case that in boys, the association between a risk factor (e.g., low birthweight) and a particular outcome (e.g., learning problems, ADHD) is stronger than it is in girls. Here, sex of the child moderates the risk relationship. A key statistical issue is that statistical power is not large to detect significant differences in measures of association across subgroups unless the sample under investigation is large. In other words, nonexperimental research is typically quite underpowered to reveal moderator processes. Furthermore, such a moderator effect does not tell us why there is a stronger relationship in one group. Only further investigation, involving processes that help to explain the reasons for such a differential linkage, can address this question.

Second, mediation also involves detecting statistical interactions, but the means of revealing such interactive processes are more complex. Here, the issue at hand is gaining understanding of how a given variable influences another. For example, let us assume that, as has often been found, there is a significant (and negative) correlation between socioeconomic status (SES) and aggression, such that for youth in lower socioeconomic subgroups, levels of aggression are higher. The mediational question is as follows: “How does this association operate?” In other words, what additional variables appear to be responsible for explaining the association?

Assume that a third measured variable reflects the amount of harsh and inconsistent discipline at home. To reveal whether this parental discipline variable mediates the SES-aggression linkage, several stages of analysis are necessary. 1) First, the predictor-criterion correlation must be significant (and clinically meaningful)—that is, there is a viable association to be explained. 2) The predictor must also show a significant association with the putative mediator; thus, low SES must be associated with harsh and inconsistent parenting. 3) The potential mediator must, in turn, be associated with the criterion measure; we assume, then, that negative parental discipline is correlated with aggression. 4) Then, in the final step, it must be shown that the predictor-criterion correlation is attenuated when the mediator variable is entered into the mix. The key point is that the presence of the mediator statistically accounts for, or explains, the basic association of interest. Full mediation would imply a complete elimination of the predictor-criterion path with inclusion of the mediator variable, whereas partial mediation indicates at least some meaningful reduction of the predictor-criterion linkage. The issue of how much the relationship must be diminished in the presence of a putative mediator variable in order to establish “mediation” is a thorny and contentious one; MacKinnon et al provide for
sophisticated elaboration. Note as well that there may be many relevant mediator processes, such that chains of mediator variables are involved.

Exploration of moderator and mediator variables is now standard fare for research in many fields of inquiry, including psychopathology. Kraemer et al. provide an expert guide to moderator and mediator variables and their role in helping to establish key conceptual and theoretical linkages in risk research. Only relatively recently, however, have these concepts been systematically adapted and applied to clinical trials methodologies.

**Application to Randomized Clinical Trials**

In a groundbreaking article, Kraemer et al. explained that when the moderator and mediator concepts are applied to treatment studies, several additional guidelines are required. In other words, because both moderator and mediator processes involve the detection of statistical interactions, further conceptual distinctions are needed to avoid confusion.3,9

In a randomized clinical trial (sometimes termed a randomized controlled trial), the independent variable is the randomly assigned treatment comparison of interest. In simple trials there are 2 levels of this variable—active treatment versus no-treatment or placebo treatment—but in more complex trials there could be additional levels (e.g., 2 or more active treatments versus a comparison condition; unimodal versus multimodal conditions treatments). The core statistical analyses focus on the main effect of treatment condition with respect to either a primary outcome measure or several distinct outcome measures, selected a priori. In order to guard against the effects of selective attrition, the standard procedure is to perform “intention to treat” analyses, such that participants are analyzed with respect to their initially assigned treatment group for the duration of the trial. The key issue, however, is that the overall results of treatment assignment do not tell us about (a) potential subpopulations that show differential response to the treatments or (b) pathways and processes by which the treatments exert their effects.

**Moderators**

Any candidate moderator variable in a clinical trial must be measured before the random assignment to intervention condition occurs. In other words, a putative moderator reflects the preexisting status of the participants, assessed during the baseline evaluations (e.g., boys vs girls; different ethnic or racial groups; high vs low SES; comorbid condition vs no comorbid condition). A moderator relationship is established when there is a significant (and clinically meaningful) interaction between the moderator variable and the independent variable of treatment condition on the outcome of interest. For example, it is conceivable that behavioral treatments for ADHD fare better with middle class as opposed to lower class families; if such is the case, SES would be a moderator of outcome.

Because the moderator measure temporally precedes the random assignment to treatment, moderator variables are, by definition, *uncorrelated* with the treatments.3,9 Thus, when investigators examine the treatment condition x moderator interaction term, they can be assured that there is no systematic bias in interpreting this interaction, because the moderator is not influenced by the random assignment. If sample sizes are sufficiently large, it is nearly certain that the treatment conditions will be balanced across levels of the moderator variable.

The presence of a moderator x treatment condition interaction signifies that, beyond the probability set by the alpha level of the statistical analysis, there is a different pattern of treatment outcomes for different levels of the moderator. Taking the example above, those youth with ADHD in middle versus low SES groups respond differently to the treatments. Beyond the statistical significance of the interaction, investigators must probe for effect size, typically by plotting the outcome measures as a function of both treatment condition and moderator variable and using preplanned analyses to probe the direction and nature of the differential effects (for details and examples, see Holmbeck). Rarely, moderator effects may exist in the absence of a main effect of treatment assignment, if the positive effects in subgroup A are canceled out by the negative effects in subgroup B.

It is not likely that moderator processes occur in isolation. In fact, anticipating analyses described below, it is now possible to use sophisticated software to analyze for the interactive effects of several moderators as they act together.13

In all, moderator variables identify subgroups of the sample with particularly strong or weak responsiveness to the interventions being evaluated. Clinically, they signal to clinicians which particular interventions are best linked with particular subpopulations.9 Conceptually, they are useful in alerting scientists to particular processes that operate in specific subgroups to enhance or diminish treatment responsiveness. In other words, moderator
findings typically prompt a search for relevant mediator variables and processes. Moderator variables have been extended to include genotypes, which may predict differential responsiveness to medication intervention for depression. Indeed, moderators range from individual-level variables all the way to social processes (e.g., SES, ethnicity, culture).

**Mediators**

Whereas a moderator variable is a baseline characteristic, mediators are measured during the time of active intervention, in an attempt to discover the processes by which interventions exert their effects. Because they temporally follow, rather than precede, the random assignment to treatment condition, mediator variables are inherently correlated with treatment assignment, rendering them susceptible to selection effects or other processes that make causal interpretation difficult. Yet they can reveal important means by which treatments may exert their therapeutic benefits and, importantly, often suggest variables to be experimentally manipulated in the subsequent clinical trials. For example, if a certain aspect of peer relationships (e.g., peer acceptance) is changed as a function of either medication or behavioral treatment for ADHD, and this change in peer status helps to explain the ultimate outcome with respect to clinically significant improvement in overall functioning, the peer acceptance could be claimed to mediate outcome. In the next trial, random assignment could be made to interventions that directly target peer-related effects, to establish more conclusively their causal relation to core outcomes.

The first consideration has to do with the timing of measurement. Mediators should be assessed during the period of intervention, suggesting strongly that a clinical trial should feature evaluation of participants in between pretreatment and posttreatment periods. At the very least, a mediator should involve a change score from before to after treatment. When a supposed mediator variable is appraised only after treatment, it is conceivable that it is simply an alternate outcome measure rather than reflecting a process that predicts change in the outcome measure. Note that a given variable might be considered as a moderator or a mediator, depending on when it is measured. For example, SES at baseline could well be a moderator of treatment but change in SES during the period of intervention could mediate outcomes. We know, for example, that reductions in externalizing behavior may well result from improvement in a family’s social standing.

Second, in terms of statistical analysis for clinical trials, the main effect of a potential mediator variable implies full mediation of the treatment effect. For instance, if improvements in peer relations fully account for the outcome, then the peer variable can be claimed as a full mediator. More typically, there is a statistical interaction of treatment condition with the mediator variable, whereby the intervention effects are enhanced in a treatment condition relative to a control condition related to change in the mediator variable of interest. Such a situation is termed partial mediation.

Third, it is often assumed that mediation can occur only when there is a significant main effect of treatment—in other words, when there is an effect of treatment to be mediated or explained—but mediation can exist even in absence of a treatment effect. Parallel to the case with moderator processes, it is conceivable that an overall null effect of treatment could mask a pattern whereby positive treatment benefits accompany change in the mediator in one direction but negative outcomes occur when the mediator goes in the opposite direction.

**Designing Clinical Trials**

Several points made above may be helpful in the design of treatment studies. First, unless there is considerable theory and empirical evidence related to particular moderator or mediator processes, in which case the trial could be designed to test such effects a priori, intervention studies need to be designed to maximize statistical power for the main effects of interest, which are the treatment comparisons regarding core outcome variables. Moderator and mediator tests are therefore exploratory, comprising secondary analyses to generate hypotheses for the next trial in a potential series. Too often, however, exploratory analyses are viewed in pejorative terms, as “fishing” or “dredging” the data set. On the contrary, appropriate tests of moderation and mediation are hypothesis generating, probing the subgroups for whom treatments optimally work (or do not work) and the processes by which treatments exert their effects. They are thus crucial for generating theoretically relevant information from clinical trials. In other words, moderator and mediator tests can help to bridge the large gap between theory and practice.

Second, it is essential that investigators choose potential moderator variables wisely, using existing theory and existing empirical findings for their selection. There is often a balance between the typical goal of an efficacy study—to generate internally valid conclusions...
Regarding the precise effects of treatments on relevant outcomes, which suggests narrow, homogeneous samples—and the more ecologically valid objectives of effectiveness studies related to having diverse samples that yield (a) external validity and generalizability of findings and (b) tests of key moderator-defined subgroups (e.g., boys vs girls, comorbid vs noncomorbid youth, children vs adolescents, high vs low SES). To the extent that the trial is sufficiently large and the sample is sufficiently varied, multiple opportunities for tests of moderator variables can emerge, so long as the variables of interest are well thought-out.

Third, theory is even more important for the selection of potential mediator variables. Does it seem likely from existing research and conceptual models that a particular cognitive, affective, familial, or peer-related process is likely to “drive” successful treatment for a given condition? Certainly in the realm of conduct problems and aggression, family discipline processes and peer rejection are linked to outcome. Thus, in an excellent example of mediational research, improvements in both caregiver discipline and enhancements of children’s social skills and peer status have been shown to mediate the success of behavioral interventions in terms of reducing aggression and delinquency. In all, investigators should blend the best features of clinical trials with the best aspects of prospective, longitudinal research to achieve the dual goals of internal validity and explanation.

Relevance to ADHD

Until relatively recently, very little treatment research on ADHD has been able to provide tests of moderator and mediator variables, largely because of the relatively small size of treatment samples and the lack of knowledge of how to design trials that could optimize conceptual yield. Furthermore, the few studies that have attempted to link process with outcome (for example, alterations in family discipline style linked to child improvement) have typically examined change in the alleged mediator process and change in the outcome variables but without explicit tests of mediation—that is, without probing whether alterations in the former domain actually explain improvements in the outcome.

The Multimodal Treatment Study of Children With ADHD (MTA Study) provided a multi-center, experimentally rigorous, large-sample means of examining (a) the relative and combined benefits of medication management and intensive behavior therapy for the long-term treatment of ADHD and (b) the impact of important moderator and mediator variables. Because the chief statistical and design consultant to the MTA was Helena Kramer, this investigation provided a pivotal test of her working model of moderator and mediator processes in the context of a large clinical trial.

The design and rationale of the MTA have been provided elsewhere. Briefly, 579 children with ADHD-Combined type, aged 7.0–9.9 years at baseline, were randomly assigned to 14 months of intervention in 1 of 4 treatment arms, each of which constituted a manualized yet flexible strategy: (a) medication management (MedMgt), in which a 28-day, double-blind, placebo-controlled stimulant trial was followed by 13 months of regular, careful monitoring of medication; (b) intensive behavior therapy (Beh), blending 35 parent training sessions, biweekly consultation with the child’s teacher, an 8-week summer treatment program, and 3 months of classroom aide support; (c) combined treatment (Comb), a multimodal combination of (a) and (b); or (d) community comparison (CC), a “treatment-as-usual” condition in which families were referred to community providers and obtained treatment of their own choosing. Quite intentionally, the sample selection procedures included common comorbidities related to ADHD and prioritized an ethnically and socioeconomically diverse sample, with a large subgroup of African-American families and a viable subsample of Latino families, as well as a 4:1 boy/girl ratio. Hence, such variables were available as potential moderators. A priori outcome measures spanned 6 domains: 3 symptom clusters (ADHD, disruptive, and internalizing) and 3 areas of functional impairment (parent-child interactions, peer relations, and academic achievement). These were assessed before, during, and after the 14-month period of active intervention, meaning that the key areas of peer relations and family interactions could be examined as mediators of symptom-related outcomes.

In all, whereas the MTA was designed to yield maximum statistical power to detect differences between the 4 treatment conditions—constituting, as such, a tightly designed efficacy study—it also included a wide range of sample characteristics, its treatments were conducted in naturalistic settings (e.g., schools), and it incorporated measures of key processes that might explain outcome. It therefore incorporated several features of an effectiveness study as well, with the intentional goal of examining key moderator variables and mediator processes.

Core findings were reported in MTA Cooperative Group. For the outcome domain of ADHD
symptomatology, the MedMgt and Comb treatments proved superior to Beh and CC. For other outcome domains, however, the most consistent finding was that Comb yielded significantly more improvement than CC; differences between the 3 active treatment arms were less consistently revealed. Furthermore, when composite outcome measures were explored in secondary analyses of the data, it was found that Comb was superior to MedMgt, with a small effect size, that MegMgt was superior to Beh, and that Beh was superior to CC, even though two-thirds of the families in this condition opted for stimulant medication from community providers. Indeed, in Swanson et al., outcomes were appraised via a categorical indicator of “excellent response,” which featured scores of less than “1” (i.e., “just a little” of the symptoms) on the 0–3 metric of an ADHD rating scale as indicative of such response, reflecting nearly normalized functioning. Across the 4 treatments, 68% of Comb met this criterion, compared with 56% of MedMgt, 34% of Beh, and 25% of CC. Yet longer-term follow-up of the MTA sample has revealed that the initial treatment group differences, evident immediately after intervention, dissipate over time, suggesting that the initial superiority of the MTA medication algorithm was not sustained in the absence of continuing, intensive intervention.

**Moderators of Outcome**

MTA Cooperative Group provided initial results of moderator tests (Table I). Neither sex (80% boys vs 20% girls) nor comorbidity with oppositional defiant disorder or conduct disorder (34% of the sample yielded such preintervention comorbidity) significantly moderated findings. Although there is little surprise that boys and girls revealed similar patterns of treatment response, it was unexpected that the aggressive/externalizing subset of the sample did not differ from the nonaggressive subset with respect to treatment response. Given the multiple problems related to associated ADHD and aggression, it might have been predicted that multimodal treatment (Comb) would have been optimal for this subgroup.

Furthermore, use of medication before to the trial (31% of the sample had such prior medication use) did not yield much in the way of moderation, with one exception: For those children assigned to Beh who had received ADHD-related medication before the study, half required remedication before the 14 months of behavioral intervention had ended. Yet the comparable statistic was that fewer than 15% of those who had never received previous pharmacologic treatment required medication during the trial. The suggestion is that once a young child with ADHD receives medication, it will be difficult to take him or her off and maintain optimal behavior.

Initial comorbidity with an anxiety disorder served as a clear moderator of treatment response. Whereas the 66% of the MTA sample without anxiety at baseline displayed a response to treatment that was close to that of the overall sample, the 34% with comorbid anxiety showed a relatively better response to the behavioral aspects of the MTA treatments. Specifically, with respect to parent-reported outcomes reflecting ADHD and internalizing symptomatology, after treatment, those ADHD-anxious children who received Beh performed just as well as did those receiving MedMgt, and Comb yielded an even better response for this comorbid subgroup. This moderator finding has important treatment implications; it may well be that, for this highly anxious subgroup, clinicians should consider beginning treatment without medication but with behavior therapy.

Jensen et al. and March et al. extended these moderator findings by considering additional comorbidity. In brief, children with ADHD plus anxiety but without oppositional defiant or conduct disorder showed markedly preferential response to Beh, whereas the subgroup with ADHD, anxiety, and disruptive behavior disorders responded optimally to Comb. It is therefore crucial to appraise single versus double comorbidities.

The MTA team also found that baseline poverty was a significant moderator variable. Specifically, 19% of the sample began the study receiving some form of public assistance. For them, 2 moderator effects emerged. First, this subgroup showed improvements in teacher-reported social skills only with Comb. Thus, for the most socioeconomically disadvantaged of the MTA participants, it took a combined medication-behavior therapy intervention (i.e., Comb) to influence arguably the most important outcome of all—social skills and peer relationships. Second, MedMgt led to a decrease in parent-reported closeness with the child across the period of active intervention for this public assistance subsample, but the other 3 treatment conditions did not yield such a decline. It may be that a child’s symptom improvements induced by pharmacologic intervention within impoverished families allow some disengagement from caregivers with the child, so that attention can be directed to other problematic areas of the family’s functioning. A possible clinical implication is that low...
SES families may show problems over a longer term if medication is the sole intervention.

What of racial background? Arnold et al. found that children from African-American backgrounds were particularly likely to benefit from Comb, even controlling for other relevant background factors. It will take additional research, with oversampling of diverse subpopulations, to replicate such results.

In an innovative report examining the effects of multiple moderator variables, Owens et al. first reiterated that, for the symptoms of ADHD and disruptive behavior considered over the period of active intervention, Comb and MedMgt led to higher rates of “excellent response” than did Beh or CC, when considering the sample as a whole. The severity of the child’s baseline ADHD symptoms served as an initial moderator, such that those youth with the most severe levels of ADHD before treatment showed under a 50% chance of excellent response with the MTA medication procedures whereas those with less severe ADHD had a greater than 70% chance of such response. Next, we examined the self-reported depressive symptoms of the primary caregiver, who was in nearly all cases the mother. This variable further moderated outcome: Those children with mothers reporting at least mild depression before the intervention began revealed a further decline in

Table I. Summary of Moderator and Mediator Findings from the MTA

<table>
<thead>
<tr>
<th>Moderator Tested</th>
<th>Significant?</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex of child</td>
<td>No</td>
<td>Boys and girls responded similarly</td>
</tr>
<tr>
<td>Comorbid ODD/CD</td>
<td>No</td>
<td>Children with and without this comorbidity responded similarly</td>
</tr>
<tr>
<td>Prior stimulant treatment</td>
<td>No</td>
<td>Children with and without prior medication histories responded similarly—except with respect to the likelihood, within Beh, that prior medication status predicts the inability to stay off medication for the full 14-month treatment period</td>
</tr>
<tr>
<td>Comorbid anxiety disorder</td>
<td>Yes</td>
<td>Children with anxiety disorders showed relatively stronger response to behavioral intervention, for parent-reported outcomes related to ADHD and internalizing symptoms, than did those without this comorbidity. Patterns of double comorbidity are also salient (ADHD + anxiety without ODD/CD predicts responsiveness to Beh; ADHD + anxiety + ODD/CD predicts responsiveness to Comb)</td>
</tr>
<tr>
<td>Family public assistance</td>
<td>Yes</td>
<td>Children of low-income families (a) showed relatively stronger response to Comb regarding teacher-rated social skills and (b) relatively less positive bonding with their parents if they received MedMgt, than did those with higher incomes</td>
</tr>
<tr>
<td>Ethnicity/race</td>
<td>Yes</td>
<td>African-American children showed relatively better response to Comb than did White children</td>
</tr>
<tr>
<td>Severity of ADHD</td>
<td>Yes</td>
<td>Children with more severe initial ADHD showed relatively worse response to MedMgt or Comb than children with less severe ADHD</td>
</tr>
<tr>
<td>Parental depressive symptomatology</td>
<td>Yes</td>
<td>Children of parents with at least mild depressive symptomatology showed relatively worse response to MedMgt or Comb than children of parents without such parental depression</td>
</tr>
<tr>
<td>Child IQ</td>
<td>Yes</td>
<td>Children with IQ scores less than 100 showed relatively worse response to MedMgt or Comb than children with IQ scores of at least 100, if they were also severe in ADHD symptomatology and had parents with at least mild depressive symptomatology</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mediator Tested</th>
<th>Significant?</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attendance</td>
<td>No</td>
<td>Not a significant mediator for Beh treatment</td>
</tr>
<tr>
<td>Attendance</td>
<td>Yes</td>
<td>A highly significant mediator for MedMgt (that is, families with high attendance showed far greater treatment response than those with low attendance for doctor’s visits</td>
</tr>
<tr>
<td>Medication use in community</td>
<td>Yes</td>
<td>Children in CC who received medication from practitioners in the community showed better ADHD-related outcome than did those who did not receive medication—but worse outcome than those in MedMgt or Comb</td>
</tr>
<tr>
<td>Parental discipline: Negative/Ineffective</td>
<td>Yes</td>
<td>Families with the greatest improvements in Negative/Ineffective discipline showed, if in Comb, relatively greater improvement, with respect to social skills and disruptive behavior in the classroom, than did those without such improved discipline or those in the other treatment conditions. The subgroup of Comb families with the greatest improvements in such discipline had normalized levels of classroom disruptive behavior by the end of active intervention</td>
</tr>
</tbody>
</table>
excellent response to the medication aspects of the treatments. Finally, the child’s baseline IQ entered the equation: For children with severe ADHD, caregivers with at least mild depression, and IQ scores below 100, there was only a 10% chance of excellent response to Comb or MedMgt. These results are sobering, as they reveal a relative failure of the intensive MTA treatment algorithms to help those children in the study who were most in need of intervention. The field has a long road to travel before the most at-risk children and families coping with ADHD are guaranteed adequate success rates, even with carefully managed and frequently monitored medication procedures.

**Mediators of outcome**

The MTA has examined several levels of mediator processes, yielding several intriguing findings (Table). First, at a basic level, it would seem intuitive that success in clinical behavior therapy depends on whether parents participate in the parent management training that is offered, teachers participate in their consultations, and children participate in the academic and social skills offered via summer treatment programs. Furthermore, it would seem necessary that, for pharmacotherapy, families attend doctor’s visits. Hence, as an initial mediator test, the MTA investigators made a priori judgments as to the levels of attendance for parent management, school consultation, and summer treatment programs and for medication visits that would be necessary for adequate response. Families were then classified dichotomously as having surpassed such thresholds or not, with such attendance examined as a potential mediator of outcome. As shown in MTA Cooperative Group, attendance levels did not influence outcome for Beh, but attendance at the pharmacotherapy sessions markedly enhanced outcome for the MedMgt and Comb. The result becomes understandable when one realizes that if the family did not attend the medication visit, they could not obtain their prescription, and the child’s behavior suffered accordingly. Thus, the initial mediator finding from the MTA was that attendance matters considerably for medication treatment but not for behavior therapy.

It may well be, of course, that a different variable (e.g., amount of actual adherence to the principles taught in the parent management or teacher consultation) would mediate outcome for behavior therapy. Thus, beyond mere attendance, actually practicing the behavioral procedures and strategies may be the “active ingredient.”

A second mediator variable was derived from a close examination of the CC condition. For this group, which received “treatment as usual,” the MTA investigators wished to determine what kinds of community intervention families were selecting. To ascertain this important variable, staff contacted families by telephone on a quarterly basis, inquiring as to the types of services the parents and child had received. From this structured telephone interview, it was discovered that 68% of the CC group obtained stimulant medication treatment for their child. Here, then, is an important potential mediator: The use of medication treatment during the active intervention period. As with all potential mediators, it is subject to selection bias: Perhaps only the most difficult children in this condition were prescribed medication; alternatively, perhaps only the most affluent or motivated parents obtained pharmacologic intervention. It is still possible, however, that this putative mediator could influence outcome.

When statistical and graphical analysis was performed, the emergent finding was that medication treatment in the CC group did yield superior outcome to the absence of medication treatment in this group. Even so, the level of posttreatment symptomatology in the medicated CC subgroup was significantly worse than that in participants receiving the MTA-delivered medication management condition. In other words, the use of stimulant medication was not identical in the CC condition as opposed to MedMgt. Specifically, in MedMgt, there was an initial 28-day titration followed by monthly maintenance sessions of 30-minute duration. Doctors prescribed t.i.d. doses (the recent, longer-acting forms of stimulants were not yet available at the time that the MTA occurred), prescribed for weekends, and contacted the child’s teacher for feedback before each monthly medication visit.

On the other hand, interview data with parents revealed that, for CC families who obtained medication for their child, the schedule of pediatric visits was once every 5–6 months, for under 10 minutes; the physicians did not prescribe an afternoon dose; and no pediatrician had, to parents’ knowledge, ever contacted the child’s teacher. In fact, the final medication dosages in MedMgt were far higher (mean methylphenidate dose of 37 mg/day) than in CC (mean dose of 22 mg/day). Each of these variables is a candidate for explaining the greater efficacy of medication as carefully prescribed and monitored in the MTA MedMgt condition than in typical pediatric office. An essential message here is that how medication is prescribed
and monitored can make a clinically significant difference in outcome.

Third, a different kind of mediator variable was examined by Hinshaw et al., who explicitly hypothesized that changes in family discipline practices would mediate school-based outcomes, particularly for families assigned to Beh and Comb. That is, the prediction was that despite the moderate to strong heritability of ADHD, along with its strong neurobiological origins, psychosocial processes in the family could still influence outcome. In other words, the prediction was that the effects of Beh and Comb on disruptive behavior and social skills—measured at school, to avoid confounding measurement of parenting with assessment of measurement of relevant outcomes—would be mediated, or explained, by improvements in family discipline practices. Contemporary theory on the importance of family context for shaping the ultimate course of ADHD helped to generate such hypotheses.35

Family discipline was assessed via parental self-report (more recently, objective measures of parent-child interactions have emerged with clear treatment effects favoring Beh and Comb; these objective indicators could serve as mediators as well). Three parenting factors emerged from a factor analysis of our parental discipline scales: Positive Parenting; Monitoring, reflecting the family’s knowledge of daily activities of the child; and Negative/Ineffective Discipline, which covered both harsh parenting practices and noteworthy inconsistencies in parenting, like promising rewards but failing to deliver them.

We followed the necessary steps for establishing mediator relations. First, treatment differences emerged for the outcomes of negative/ineffective parenting and social skills.15 Second, treatment differences emerged for the putative mediator of parenting, particularly Negative/Ineffective Discipline; here, Comb and Beh showed the strongest reductions. Third, these parenting practices influenced social skills at school and, to a lesser extent, teacher reports of disruptive behavior; thus, the potential mediator was influential regarding key outcomes. Fourth, the crucial interactions emerged, such that treatment condition interacted significantly with discipline style in terms of affecting outcomes.

This final step is important to examine in more detail. For social skills, a clear pattern of mediation was found, such that the superiority of Comb for this critical outcome was mediated by parental reductions of Negative/Ineffective Discipline. In other words, for those families with large improvements in such discipline during the course of the study, the outcome of social skills at school showed important improvements, related mainly to those families receiving Comb. Thus, reductions of such harsh and inconsistent discipline mediated, or explained, the treatment effect on social skills, with the strongest benefits noted for Comb.

For the outcome measure of teacher-rated disruptive behavior (ADHD plus aggressive symptoms), the pattern of mediation was not as clear-cut because specific treatment contrasts did not always yield a significant effect of treatment condition on this outcome. Importantly, however, for the subgroup of families receiving Comb whose negative/ineffective Discipline scores improved substantially, mean scores by the end of active treatment were at the level of national norms for the rating instrument used.17 In other words, the only subgroup of children within the entire MTA trial who showed normalized posttreatment levels of disruptive behavior at school were those (a) who obtained Comb, the multimodal combination of medication plus intensive behavior therapy and (b) whose parents showed substantial improvements in the types of negative, harsh, inconsistent, and ineffective discipline practices that appear to promote disruptive behavior and exacerbate externalizing problems.36 Thus, even for a condition with clear neurobiological roots and substantial heritability, like ADHD, parental socialization is associated with important effects on important clinical outcomes. Even though problematic parenting does not appear to be a causal factor for ADHD, changes in parenting practices may be highly related to long-term course of this condition. Subsequent trials should emphasize randomized assignment to treatment conditions that directly attempt to alter family discipline.

Summary

The MTA has yielded an unprecedented degree of information regarding moderators and mediators of behavioral and medication treatment effects for children with ADHD. Before the MTA, sample sizes were too small for adequate tests of such variables, and sound knowledge as to how to think about examination of these processes was lacking. Several findings are salient regarding treatment moderation: Whereas boys and girls with ADHD appear to show nearly identical patterns of response to treatment, and whereas comorbidity with externalizing behaviors does not moderate treatment response, comorbidity with anxiety disorders predicts a preferential response to behavioral treatment, suggesting that for children with this comorbidity, clinicians should...
consider behavioral intervention as a first-line treatment strategy. Furthermore, for families receiving public assistance, combined medication-behavioral treatment yields important improvement in social skills and may mitigate the tendency for medication alone to be linked to a lack of closeness between children and parents. It is suggested, as well, that African-American children show a preferential response to combined medication-behavioral treatment strategies. When multiple moderators are considered together, the distressing result is that, even with well-delivered and carefully monitored medication procedures, children with the most severe levels of ADHD, children whose primary caregivers have depressive symptoms, and children with subaverage IQ scores show decreased likelihood of excellent response. Thus, for the children with ADHD most in need of treatment, current interventions are falling short.

With respect to mediators, the following points are salient. First, somewhat surprisingly, attendance at psychosocial aspects of treatment does not help to explain outcome, although more sophisticated indicators of treatment performance may be salient. Second, the success of medication treatment appears to depend heavily on adequate monitoring and dosing. That is, when treated pharmacologically with infrequent monitoring and a lack of afternoon dosing, children with ADHD-Combined type show far less improvement than do those with well-delivered and monitored medication procedures (i.e., MedMgt and Comb conditions). Third, clinically significant improvements in social skills and clinically significant reductions in disruptive behavior at school are explained, in part, by enhancement of parenting style—in particular, by reductions in negative/ineffective discipline—particularly for families receiving Comb, the multimodal intervention. Thus, even for a condition, like ADHD, with clear neurobiological underpinnings, changes in family environment appear crucial for explaining therapeutic change. This latter finding should help to deliver clear messages to parents and caregivers of youth with ADHD who enter treatment: Although such caregivers are not to blame for having caused their child’s ADHD, they are clearly responsible for changing disciplinary practices in order to produce important treatment-related gains.

Extensions and Future Directions
Many commentators on child psychotherapy and psychopharmacology have called for greater attention to be paid to understanding which treatments work best for which children and families—and how and why treatments exert their effects. Thus, mediator and moderator processes are now a major component of most important clinical trials.

Despite the growing interest in such variables and processes, what is known far outweighs what is known. For example, with respect to moderator processes, almost no knowledge now exists regarding the factors that predict positive outcome for children with the Inattentive type of ADHD. Additionally, whereas genotypes are increasingly examined as potential moderators of treatment response for adult conditions (e.g., mood disorders), the field does not yet have validated evidence in this regard for ADHD. Furthermore, an extremely important therapeutic goal will be to discover whether children from different ethnic groups, cultures, and nations show differential responsiveness to the empirically established interventions for ADHD. Such macro-level moderators are extremely important to examine in future treatment research.

As for mediators, vast realms of knowledge are still underexplored. At a pharmacologic level, what are the specific, dynamic processes at the neural level that “carry” the effects of medication treatments? Several processes related to dopaminergic (and noradrenergic) neurotransmission have been implicated, but the precise mechanisms are not yet fully known. Regarding psychosocial processes, it is inaccurate to claim that the genetic/biological underpinnings of ADHD negate any role for parenting (or schooling) variables in shaping treatment response. Even so, the MTA finding that reductions in negative/ineffective discipline appear to explain treatment response mandates replication; and the kinds of school environments that promote optimal therapeutic gains are not yet fully elucidated.

Future research on interventions for ADHD should incorporate the following: 1) an appreciation of conceptual and theoretical models regarding the processes underlying both the genesis and maintenance of this condition, so that the appropriate variables can be measured and the optimal interventions can be tested; 2) sound research designs that afford unambiguous interpretation of main effects of various treatments; and 3) careful measurement of relevant moderator and mediator variables, in order that particularly responsive subgroups and particularly important treatment processes become known. In this way, conceptual gains as well as clinical knowledge can emerge from clinical trials related to ADHD, potentially enhancing the further specificity and power of later clinical trials. Indeed, it
is noteworthy that very few treatment trials in all of child and adolescent psychiatry use mediator analyses to understand processes of change.\textsuperscript{38} Furthermore, understanding of factors that promote long-term change is essential, given that ADHD is not a short-lasting problem but rather a chronic condition. Overall, understanding of moderator-defined subgroups and mediator-related processes is not a luxury but rather a necessity if progress regarding treatment for ADHD, as well as all other child/adolescent disorders, is to proceed with precision, rigor, and clinical and conceptual elucidation.

Acknowledgments
Supported in part by NIMH grants R01 45064, R01 50461, and N01 12009. I express my appreciation to the many children and families who have participated in the clinical studies described herein.

Received November 28, 2005; accepted April 1, 2006

References


