Abstract: With recent increases in the use of positive approaches to treatment for individuals with developmental disabilities, it seems appropriate to review the variables that have been found to influence the acceptability of various treatments. Programmatic treatments for problematic behaviors that incorporate primarily positive (reinforcement) components rather than negative (punishment) components may still be susceptible to variables found to influence the acceptability of treatments. Although more positive reinforcement based approaches are certainly preferred, the need to consider the right to effective treatment is also an important component of any intervention for problematic behavior. To continually assure the right to effective treatment, the examination of variables affecting the acceptance of treatments continues to be an important area of research. This paper reviews the instruments that have been used to evaluate the acceptability of treatments as well as the variables that have shown demonstrated influence on the acceptability of treatments for problematic behavior.

Research in the area of treatment acceptability is relatively new to the area of education and psychology with a noticeable increase in treatment acceptability studies occurring during the 1980’s (Elliott, 1988). Recognition of the need for research in the area of social validity became more evident with an elaboration of the topic in an article by Baer, Wolf, and Risley (1968). Prior to this publication, the primary emphasis with regard to educational and behavioral treatments had been treatment efficacy (Witt, 1986). Wolf (1978) stated that social validity refers to decisions based on the social importance of interventions on three related levels. First, objectives of the treatment must be socially significant in which the desired outcome is both important and pertinent. Second, treatment procedures must be deemed as socially appropriate in that they are generally considered acceptable by society. Third, effects of the treatment should have clinical significance. Kazdin (1980) noted that social validity research has mainly focused on the appropriateness of treatment procedures, an area of research also known as treatment acceptability. The term treatment acceptability, a component of social validity, was defined by Kazdin as judgments of treatments by actual or potential consumers of the treatments, such as nonprofessionals, clients, laypersons, and others. For the purposes of this review, treatment acceptability will be defined according to the definition proposed by Kazdin and the terms treatment and intervention will be used interchangeably.

Treatment Acceptability Measurement Instruments

Two of the most frequently used instruments for measuring treatment acceptability are the Treatment Evaluation Inventory (TEI; Kazdin, 1980) and the Intervention Rating Profile (IRP; Witt & Elliott, 1985). Both of these instruments have been modified and revised in multiple ways by researchers in attempts to address target raters more effectively or to reduce the administration time of the instruments. Because modifications to these instruments have been implemented in various acceptability studies, a review of these two instruments and their subsequent modifications is provided (refer to Table 1).

Treatment Evaluation Inventory. The TEI was originally designed to assess treatments recommended for children with behavior disorders (Kazdin, 1980). It consists of 15 items that ask direct questions regarding specific treatment procedures. Ratings are made on a seven point Likert-type scale, with descriptive anchor points that vary depending on the...
### TABLE 1

**Treatment Acceptability Instruments**

<table>
<thead>
<tr>
<th>Instrument Name</th>
<th>Type of Instrument</th>
<th>Validity</th>
<th>Reliability</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment Evaluation Inventory (TEI; Kazdin, 1980)</td>
<td>Consists of 15 items, uses a seven point Likert-type scale, descriptive anchor points vary depending on question</td>
<td>Item loadings from .61 to .95 on a unitary factor accounting for 51.4% of variance</td>
<td>Internal consistency ranged from .35 to .96</td>
<td>Considered the first and one of the most frequently used instruments</td>
</tr>
<tr>
<td>Treatment Acceptability Rating Form (TARF; Reimers &amp; Wacker, 1988)</td>
<td>Consists of 15 items, uses a seven point Likert-type scale, varying descriptive anchor points</td>
<td>Internal consistency ranged from .80 to .91</td>
<td>Revised version of the TEI, developed by Reimers and Wacker, incorporated factors of effectiveness, and cost of treatment</td>
<td></td>
</tr>
<tr>
<td>Treatment Evaluation Inventory Short Form (TEI-SF; Kelly et al., 1989)</td>
<td>Consists of nine items rated using a five point Likert-type scale with fixed anchor points</td>
<td>Internal consistency of .85</td>
<td>Revised version of the TEI-SF, requires less time to complete than the TEI-SF</td>
<td></td>
</tr>
<tr>
<td>Treatment Acceptability Rating Form-Revised (TARF-R; Reimers et al., 1991)</td>
<td>Consists of 20 items rated using a seven point Likert-type scale, varying descriptive anchor points for each item</td>
<td>Internal consistency of .92</td>
<td>Revised version of the TARF, includes items addressing problem severity and understanding of treatment</td>
<td></td>
</tr>
<tr>
<td>Intervention Rating Profile (IRP; Witt &amp; Elliott, 1985)</td>
<td>Consists of 20 items evaluated by a six point Likert-type scale</td>
<td>One primary factor (41% of the variance) and four secondary factors</td>
<td>Internal consistency of .89</td>
<td>Designed to evaluate educational treatments</td>
</tr>
<tr>
<td>Intervention Rating Profile-15 (IRP-15; Martens et al., 1985)</td>
<td>Consists of 15 items, uses a six point Likert-type scale</td>
<td>Principal component analysis resulted in a unitary factor</td>
<td>Internal consistency of .98</td>
<td>Modified the IRP to increase internal consistency</td>
</tr>
<tr>
<td>Children’s Intervention Rating Profile (CIRP; Witt &amp; Elliott, 1985)</td>
<td>Consists of seven questions evaluated using a seven point Likert-type scale</td>
<td>Principal component analysis resulted in a unitary factor</td>
<td>Internal consistency ranged from .75 to .89</td>
<td>Modified the IRP for children, items written at fifth grade reading level</td>
</tr>
<tr>
<td>Behavior Intervention Rating Scale (VonBrock &amp; Elliott, 1987)</td>
<td>Consists of 24 items, uses a six point Likert-type scale</td>
<td>Internal consistency of .97</td>
<td>Modified version of the IRP-15</td>
<td></td>
</tr>
<tr>
<td>Abbreviated Acceptability Rating Profile (AARP; Tarnowski &amp; Simonian, 1992)</td>
<td>Consists of eight items rated using a six point Likert-type scale</td>
<td>Unitary factor accounted for 84.9% of the variance with item loadings ranging from .89 to .96</td>
<td>Internal consistency of .98</td>
<td>Modified the IRP-15 to reduce completion time and improve readability</td>
</tr>
</tbody>
</table>
question. Total scores are obtained by summing all items with higher summed total scores indicating greater levels of treatment acceptability. A principal component analysis of the TEI rotated to varimax criterion resulted in item loadings from .61 to .95 on a unitary factor accounting for 51.4% of the variance (Tarnowski & Simonian, 1992). The TEI is the most frequently used measure of treatment acceptability, but it has been used very little in actual clinical intervention studies to measure acceptability of behavioral interventions with children and families. Modifications to the TEI have been developed in order to decrease administration time and to make it more comprehensive.

**Treatment Evaluation Inventory-Short Form.** The TEI was abridged and published as the Treatment Evaluation Inventory-Short Form (TEI-SF; Kelley, Heffer, Gresham, & Elliott, 1989). This revised form used to measure acceptability of treatments for children with behavior disorders was reduced to nine items to make it more expedient to complete. Items are rated using a five point Likert-scale with fixed, anchored points. Descriptions on the Likert-type scale range from 1 (strongly disagree) to 5 (strongly agree). Total scores are obtained by summing all items with higher summed scores indicating greater levels of acceptability. The internal consistency of this instrument was reported to be .85 (Kelley et al.).

**Treatment Acceptability Rating Form.** The Treatment Acceptability Rating Form (TARF; Reimers & Wacker, 1988) was devised from Kazdin’s (1980) original measure of treatment acceptability to measure parent’s acceptability of treatments devised within a clinical setting. It consists of 15 items on a seven point Likert-type scale. The Likert-type scale range from 1 (strongly disagree) to 6 (strongly agree). Total scores are obtained by summing all items with higher summed scores indicating greater levels of acceptability. Principal component analysis with varimax rotation revealed a primary factor accounting for 41% of the variance and four secondary factors (Tarnowski & Simonian, 1992). The internal consistency of this instrument was reported to be .92.

**Intervention Rating Profile.** The IRP was developed to extend research in treatment acceptability to educational treatments and make practitioners more aware of interventions considered acceptable by teachers. It consists of 20 items that are rated on a six-point Likert-type scale. The Likert-type rating scale ranges from 1 (strongly disagree) to 6 (strongly agree). Total scores are obtained by summing all items with higher summed scores indicating greater levels of acceptability. Principal component analysis with varimax rotation indicated two primary factors accounting for 41% of the variance and four secondary factors (Tarnowski & Simonian, 1992). The internal consistency of this instrument was reported to be .92.

**Intervention Rating Profile – 15.** Modifications were made to the IRP to shorten the instrument and to increase item loading on a single factor. This modified version of the IRP was published as the Intervention Rating Profile – 15 and used to measure educational interventions (IRP-15; Martens, Witt, Elliott, & Darveaux, 1985). Modifications to improve the internal consistency of the instrument included generating eight new items and removing other items resulting in a total of 15 items. Items are rated using a six point Likert-type rating scale with ranges from 1 (strongly disagree) to 6 (strongly agree). Total scores are obtained by summing all items with higher summed scores indicating greater levels of acceptability. The internal consistency of this instrument was reported to be .98.

**Children’s Intervention Rating Profile.** The Children’s Intervention Rating Profile (CIRP; Witt & Elliott, 1985) was a modification of the IRP designed to assess treatment acceptability of educational interventions by children. The instrument consists of seven questions related to the perceived fairness and expected effective-
ness of a treatment. Questions on this instrument were written at a fifth-grade reading level. The items are rated using a seven point Likert-type rating scale. Total scores are obtained by summing all items with higher summed scores representing greater levels of acceptability. The internal consistency of this instrument ranged from .75 to .89.

Behavior Intervention Rating Scale. The Behavior Intervention Rating Scale (BIRS; Brock & Elliott, 1987) was developed to create a new instrument for measuring treatment acceptability in school settings. The BIRS is a modification of the IRP-15, which added nine items for a total of 24 questions included on the instrument. Items are rated using a six point Likert-type rating scale that ranges from 1 (strongly disagree) to 6 (strongly agree). Total scores are obtained by summing all items with higher summed scores indicating greater levels of acceptability. The internal consistency of this instrument was reported to be .97.

Abbreviated Acceptability Rating Profile. An additional abbreviated form of the IRP-15 was developed to measure educational interventions and published as the Abbreviated Acceptability Rating Profile (AARP; Tarnowski & Simonian, 1992). They developed the AARP through modification of the IRP-15 by eliminating seven items and maintaining eight items that were reworded to improve readability. Items are rated using a six point Likert-type rating scale that ranges from 1 (strongly disagree) to 6 (strongly agree). Total scores are obtained by summing all items with higher summed scores indicating greater levels of acceptability. The internal consistency of this instrument was reported to be .98.

Psychometric properties of the AARP were initially evaluated using a sample of 60 parents seen for routine pediatric outpatient visits at a large urban hospital. The initial sample when subjected to a principal component analysis resulted in all items loading on a unitary factor (Acceptability) that accounted for 84.9% of the variance with item loadings ranging from .89 to .96. The initial sample was cross-validated in a second independent sample of 80 mothers by subjecting the data to identical analysis and resulting in item loadings ranging from .89 to .98 on a single factor that accounted for 90.3% of the variance. The AARP resulted in reduced completion time to approximately one-half the ten minutes required to complete the IRP-15. Readability analyses conducted using the Harris-Jacobson Wide Range Readability Formula (Harris & Jacobson, 1982) resulted in indices of 5.0 for the AARP compared to 7.9 for the IRP-15.

Comparison of Treatment Acceptability Measures

Finn and Sladeczek (2001) critiqued nine treatment acceptability measures which were the: (a) TEI, (b) TEI-SF, (c) TARF, (d) TARF-R, (e) IRP, (f) IRP for teachers, (g) AARP, (h) CIRP, and (i) BIRS. These nine measures were selected after an extensive database search of studies investigating behavioral intervention acceptability. The evaluation covered eight areas of treatment acceptability as follows: (a) definition of treatment acceptability; (b) content and purpose; (c) test reliability; (d) test validity; (e) statistical analysis; (f) sample characteristics; (g) scoring procedures; and (h) uses of the measure in research and practice. The evaluation indicated no single measure of treatment acceptability to be more comprehensive than any other. Conversely, in comparison to the IRP-15, Tarnowski and Simonian (1992) indicated that the AARP was more simplistic, easier to read, and required less administration time. In addition, Kelley et al. (1989) considered the AARP to be an improvement over the IRP-15 as well as other measures of acceptability.

Method

Although treatment acceptability research methodology has primarily consisted of analogue research, several variations of this methodology have been incorporated to extend the ecological validity of the findings. Variations have included presentation of case descriptions of problem behavior and treatments by use of audio tape (Kazdin, 1980), written summaries (Cavell, Frentz, & Kelley, 1986), video presentations (Martens et al., 1985), and combinations of presentation formats including actual implementation of treatments (Reimers, Wacker, Cooper, & De Raad, 1992). Participants may rate the treatment acceptability of one treatment applied to several different case descriptions or rate treatment ac-
ceptability of several treatments applied to the same or several different case descriptions. While analogue research methodology has provided insight into the acceptability of several treatments when hypothetically applied to numerous different cases covering a wide range of variables, the findings may be limited in ecological validity when compared to a clinical research methodology (Miltenberger, 1990). Clinical research methodology in treatment acceptability has typically involved having clients and/or consultees rate treatment acceptability prior to, during, and/or after implementation of a treatment. While clinical research methodology to investigate treatment acceptability appears to provide more ecologically valid findings and naturalistic evaluation of treatment acceptability in relationship to other variables, the analogue method of investigating treatment acceptability provides a much larger and more representative data base more quickly especially with treatments for low incidence populations.

Previous Literature Reviews of Treatment Acceptability

Reimers, Wacker, and Koepple (1987) conducted a review of the literature on behavioral interventions and found five primary factors that were considered to affect treatment acceptability. These factors were problem severity, treatment approach, time needed to implement treatment, side effects of treatment, and cost. Treatments for more severe problem behaviors were generally considered more acceptable than treatments for less severe problem behaviors, with some evidence for increased acceptability when more restrictive treatments were matched with more severe problems and less restrictive treatments were matched to less severe problems. Reinforcement-based procedures were generally rated more acceptable than punishment-based procedures. Treatments that required less implementation time were generally rated as more acceptable in comparison to treatments requiring more implementation time. One study reported that treatments with adverse side effects were rated less acceptable than those reporting no side effects. Cost was a factor suggested by Reimers et al. to influence treatment acceptability, although no studies were reported to have examined this factor as a variable.

Miltenberger (1990) conducted a review of the literature on treatment acceptability research conducted during the 1980’s and suggested that the most acceptable treatments would be those which were least restrictive, required little time, have the fewest side effects, are least disruptive to other students, are consistent with the rater’s training or orientation, are presented with the most appropriate rationales, are considered to be necessary for behavioral improvement, and are considered to be most effective.

Elliott (1988) reviewed 20 empirical studies on treatment acceptability of behavioral interventions for school children as rated by teachers, children, and psychologists. Elliott concluded that treatment acceptability was a complex construct influenced by several variables. Additionally, Elliott indicated that educational consumers rated positive treatments as more acceptable than reductive treatments and that there was a moderate-to-strong relationship between pretreatment acceptability ratings and perceived treatment effectiveness.

The present literature review does not attempt to provide a comprehensive review of literature on treatment acceptability conducted prior to 1990 (See Elliott, 1988; Miltenberger, 1990; and Reimers et al., 1987 for detailed reviews of treatment acceptability research published prior to 1990). Representative studies conducted prior to 1990 are reviewed in brief as examples of previously established themes found within the treatment acceptability literature. These representative studies are provided as a comparison to more recent research on treatment acceptability.

Review of Recent Treatment Acceptability Research

This review focuses on recent treatment acceptability research, which was defined as studies published during the 1990’s through 2005, while comparing these recent studies with some representative studies conducted prior to 1990. This segment of the literature was chosen in order to update the most recent review of treatment acceptability, which was considered to be Miltenberger’s (1990) review. Studies were identified by conducting a search in December 2005 for the terms “treat-
ment acceptability” using the search engine EBSCO Host. Additional articles were identified through the reference section of articles reviewed. An emphasis was placed on discussing treatment acceptability research that directly involved studies focusing on treatments for problem behaviors. Studies primarily focusing on academic skill enhancement, counseling techniques, and consultation methods were excluded from this review.

The recent review involved examining the research on treatment acceptability with regard to three types of variable manipulation. Although these three types of research often overlap with multiple variable manipulations, for the purposes of this review, the specific variables primarily focused on in the study were examined along these three separate lines of research (i.e., manipulations involving treatment variables, client/case description variables, and rater variables/individuals from which treatment acceptability ratings are obtained). Further clarification of these lines of research is provided below.

**Treatment Variable Manipulation**

Research involving acceptability of interventions with the manipulation of treatment variables appears to have received the greatest attention within the literature on treatment acceptability (Kazdin, 1980; Spreat & Walsh, 1994). Prior to 1990, treatment variable manipulation primarily consisted of manipulations of different types of treatment, varying effectiveness of treatment, and different mediators of treatment. Manipulations of different types of treatment has included factors such as comparing the acceptability of treatments that are highly intrusive versus less intrusive (Kazdin), based on reinforcement techniques versus more punishment based techniques (Kazdin; Witt, Elliott, & Martens, 1984), apparent appropriateness (Cavell et al., 1986), and time required to implement treatment (Witt et al.). Research involving manipulations of the effectiveness of treatments has evaluated the acceptability of treatments by: (a) including or excluding information on treatment efficacy (Brock & Elliott, 1987); (b) providing information on the efficacy of the treatment during previous treatment (Kalfus & Burk, 1989); and (c) providing information on the efficacy of the treatment in comparison to other treatments (Clark & Elliott, 1988). Manipulations involving the mediator of treatment have compared acceptability of treatments proposed to be implemented by parents and teachers versus school psychologists (Kalfus & Burk). Manipulations involving the professional recommending a treatment have involved comparisons of treatments recommended by a teacher, school psychologist, and a pediatrician (Carter, 2005). The following sections will discuss studies that represent the different types of treatment variable manipulation conducted prior to 1990 and provide a review of more recent literature (1990–2005), which extends or complements the research conducted prior to 1990.

**Intrusiveness of treatments.** Kazdin (1980) conducted a treatment variable manipulation experiment to assess acceptability of alternative treatments for inappropriate behaviors of children. The experiment involved 68 female and 20 male undergraduates who rated acceptability of four different treatments (reinforcement, time-out from reinforcement, drug treatment, and electric shock) each of which represented increasing levels of intrusiveness, respectively. Case vignettes were presented via a cassette tape player operated by the experimenter in a separate observation room from the participant. Treatment acceptability data was obtained by using the TEI. Results indicated statistically significant differences among all treatments with the following hierarchy of acceptability respectively, reinforcement, time-out from reinforcement, drug treatment, and electric shock. This study demonstrated that highly intrusive treatments were considered less acceptable than less intrusive treatments. Spreat and Walsh (1994) extended the examination of the intrusiveness of treatment on treatment acceptability. They surveyed members of American Association of Mental Retardation (AAMR). Treatment intrusiveness variables described within the case vignette included previously used procedures, likely side effects, and different treatments. Intrusiveness factors found to influence acceptability were restrictiveness of the proposed treatments and whether other procedures had been previously tried. Jones, Eyberg, Adams, and Boggs (1998) continued the examination of intrusiveness of
treatments. They assessed the treatment acceptability of six child management techniques (positive reinforcement, response cost, differential attention, time-out, overcorrection, and spanking) using the TEI-SF with 20 mothers of children referred for treatment of disruptive behavior. Mothers rated the positive reinforcement technique as more acceptable than other techniques while spanking was rated as less acceptable than the other techniques.

Reinforcement versus punishment. The previously described study by Kazdin (1980) also revealed a relationship between treatment acceptability with reinforcement based techniques and punishment based techniques. The reinforcement procedure as described in the study represented a reinforcement based treatment and all other treatments represented punishment based techniques. The reinforcement based treatment was found to be the most acceptable of all the treatments used in the study. Burgio et al. (1995) continued the research on reinforcement techniques compared to punishment techniques. They evaluated the acceptability of two behavioral treatments (differential reinforcement and time-out) and a drug therapy regime among elderly individuals using a modified version of the TEI. Ratings demonstrated significant differences between all treatments for behavioral disturbances of geriatric individuals with differential reinforcement receiving the highest ratings followed by the time-out, and drug therapy. In addition, significantly higher acceptability ratings were given to time-out procedures described as taking place in a nursing home versus a community setting. Several other studies have examined procedures consisting of reinforcement compared to punishment techniques although this comparison was not the primary focus of the study. These studies are described under the category that appeared to be the primary focus of the research.

Apparent appropriateness of treatments. Cavell et al. (1986) used the TEI with 120 middle and high school teachers to assess the acceptability of four paradoxical treatments with different rationales and an ineffective contingency contract for a hypothetical case of school disruption and truancy. They found that continuing the ineffective contingency contract was rated as more acceptable than all paradoxical treatments regardless of the rationale provided. Betts and Remer (1993) evaluated the acceptability of paradoxical versus nonparadoxical interventions. Paradoxical interventions were defined as techniques designed to eliminate undesirable behavior by encouraging the undesirable behavior. Participants in this study were 97 undergraduates taking part in a semester long simulation of a family arguing with a rebellious adolescent daughter. Participants were grouped into “families” and participated in four role-play exercises and a family therapy session prior to evaluating either a paradoxical or nonparadoxical intervention using the TEI-SF. Findings indicated that paradoxical interventions were less acceptable than nonparadoxical interventions, although paradoxical interventions were considered to be an acceptable intervention overall.

Time required to implement treatment. Witt et al. (1984) examined the influence that varying amounts of time associated with implementing a treatment had on the acceptability of the treatment. They found that the less time required to implement a treatment, the more acceptable the treatment would be rated.

Previous treatment effectiveness. Kalfus and Burk (1989) examined the influence of previous treatment effectiveness on treatment acceptability. They administered the TEI to 105 graduate level education students and 53 undergraduate level psychology students to evaluate the acceptability of five treatment procedures for a case vignette of a child who engaged in pica. Five treatment procedures were rated in the following order of most acceptable to least acceptable: positive reinforcement, overcorrection, contingent removal of the pica item, differential attention, and time-out. Treatment history information was manipulated by providing either a long case description or a shorter case description. The absence of treatment history information increased acceptability ratings by undergraduate psychology students and decreased acceptability ratings by graduate education students. Findings from this research supported previous research by identifying positive reinforcement as the most acceptable treatment, time-out as the least acceptable treatment,
and other treatments receiving ratings falling between positive reinforcement and time-out.

Comparison of alternative effectiveness of treatments. Clark and Elliott (1988) examined the influence of manipulating the level of treatment effectiveness on treatment acceptability. They distributed a case vignette, a treatment acceptability measure, and a general knowledge of techniques questionnaire to 133 elementary school teachers from Nebraska and Louisiana. The case vignettes manipulated two potential treatments: modeling-coaching (a form of) versus overcorrection method (a form of) and two levels of outcome effectiveness (weak vs. strong therapeutic effects). Treatment acceptability was rated using the BIRS. Findings indicated a statistically significant preference for the modeling-coaching treatment when compared to the overcorrection treatment and statistically significant positive influences on acceptability by strong therapeutic outcome effects.

Mediator of treatment. Kalfus and Burk (1989), in the study previously described, evaluated the effects of different treatment mediators on treatment acceptability. In their study, treatment mediator identity was manipulated by presenting the treatment mediator as either a psychologist or as a combination of parents and teachers. The identity of the treatment mediator (parents and teachers, or psychologist) did not influence acceptability ratings.

Professional recommending treatment. Carter (2005) found that the title of the individual recommending a treatment for problem behaviors influenced the acceptability of treatments as rated by college students. The study compared the acceptability of medication, token economy with response cost, and a time-out procedure for problem behavior when recommended by a special education teacher, a school psychologist, and a pediatrician. Findings revealed significantly lower acceptability ratings for the token economy with response cost when recommended by a physician.

The treatment acceptability literature that has focused on treatment variable manipulation has determined some specific treatment variables that influence acceptability ratings. Another line of research within the treatment acceptability literature has involved manipulating variables associated with the client. Research involving client variable manipulations will now be described.

Client Variable Manipulation

Another line of research has involved the manipulation of variables associated with the client described within the case description and their impact on treatment acceptability. This line of research has consisted primarily of manipulations of: (a) the severity of client problem (Kazdin, 1980), (b) the age/gender of the client (Elliott & Fuqua, 2002; Spreat & Walsh, 1994), and (c) the diagnostic label of the client (Fairbanks & Stinnett, 1997; Stinnett, Crawford, Gillespie, Cruce, & Langford, 2001). Research prior to 1990 involving client variable manipulation consisted of manipulations of the severity of client problem (Kazdin) and indirectly evaluating treatments for several different problem behaviors such as noncompliance, self-injury, aggression, etc. (see Miltenberger, 1990 for comprehensive list).

Severity of client problem behavior. Kazdin (1980) evaluated the impact of the severity of client problem behavior on treatment acceptability. The experiment included 68 female and 26 male undergraduate participants who rated acceptability of four different treatments (reinforcement, time-out from reinforcement, drug treatment, electric shock) using the TEI. Additionally, manipulations of the clients’ problem behavior severity (moderate vs. severe) were conducted within the case vignettes. The case vignettes were presented via cassette tape operated by the experimenter in a separate observation room from the participant. Results indicated the reinforcement treatment as statistically significantly more acceptable than other treatments while electric shock was statistically significantly less acceptable than other treatments. The drug therapy and the time-out from reinforcement treatments did not differ on acceptability. The severity of the problem behavior produced a statistically significant increase in overall acceptability ratings.

Tingstrom (1990) used the IRP-15 to assess 103 teacher ratings of a time-out procedure when implemented by a teacher or a school psychologist for either mild or severe problem
behavior. Findings revealed no significant differences based on the individual described as implementing the treatment. Significantly higher ratings were provided for time-out when applied to severe problem behavior vs. mild problem behavior.

Reimers et al. (1992) conducted both analog and clinical examinations of 40 parent ratings of treatment acceptability using the TARF-R. Parents rated the acceptability of one of three treatments (positive reinforcement, time-out, or medication) for a case description of a fictional child displaying either mild or severe problem behaviors. In addition, parents were given written descriptions and provided demonstrations of positive treatment packages (e.g., verbal praise, differential reinforcement, token systems, etc.) which were recommended for implementation with their child. Parents rated the acceptability of the recommended treatment package prior to implementing the package and at 1, 3, and 6 month follow-ups. Findings of the analog evaluations revealed higher acceptability ratings for positive reinforcement and time-out when recommended for the mild problem behaviors and higher ratings for medication when recommended for the severe problem behaviors. The influence of child problem severity on acceptability for the clinical ratings was conducted by dividing the parents into two groups based on ratings obtained from a problem behavior checklist. Findings from the clinical group indicated that parents of children with less severe problems rated the positive treatment packages more acceptable than parents of children with more severe problem behavior on all ratings obtained except for the 3-month follow-up which revealed no differences in ratings. Findings demonstrated similar ratings of treatment acceptability by parents across analog and clinical contexts. Additional analyses revealed that treatment acceptability ratings may influence treatment compliance and treatment efficacy.

Age/gender of client. The impact of client age on treatment acceptability was evaluated by Elliott and Fuqua (2002) using the AARP. They evaluated the acceptability of four interventions (habit reversal, hypnosis, medication, and punishment) for treating trichotillomania. The study presented case vignettes to 239 college students in which the age of the client (8 years, 16 years, 26 years) and the severity of hair pulling (mild vs. severe) were manipulated. Results showed significant differences among the four treatments with hypnosis and habit reversal being rated as the most acceptable interventions. The age of client and severity of trichotillomania did not significantly alter the ratings.

Spreat and Walsh (1994) assessed factors associated with decisions regarding acceptability of behavior modification programs by members from differing divisions of the AAMR. Surveys were mailed to 400 members of Region IX of the AAMR. Of these surveys, 198 were returned representing approximately a 50% return rate. The survey consisted of a case vignette with manipulations occurring among nine variables with randomly assigned values. Variables manipulated for the client described in the case vignette included sex, age, level of mental retardation, restrictiveness of residence, behavior descriptors, severity of self-injury, and frequency of behavior. A modified TEI was used to rate the acceptability of each case vignette. None of the client variables were found to be statistically significant in influencing treatment acceptability.

Diagnostic label of the client. Stinnett et al. (2001) manipulated the variable of attaching a label to the client described in the case description and the impact this manipulation had on treatment acceptability. Their study also evaluated the acceptability by teachers-in-training of two treatments, Ritalin versus special education placement, for students with a label of ADHD versus those without a diagnosis. Participants were recruited from undergraduate teacher education courses at a medium-sized university in the southwest and consisted of 27 males and 117 females. Participants were predominantly Caucasian (87.5%) with 50% having graduated from a rural high school and 49% having graduated from an urban high school. Participants were presented with a case vignette with manipulations involving label (ADHD vs. no label) and treatment (Ritalin vs. special education placement). Participants self-reported high school location at graduation and rated acceptability of interventions using the IRP-15 (Martens et al., 1985). These data were analyzed with other data collected using three 3-way ANO-
VAs. The impact of labeling on the child in the case vignette with ADHD versus no label was found to have no statistical significance and no statistically significant difference was found between the two treatments (Ritalin vs. special education placement).

Miller, Manne, and Palevsky (1998) used the TEI-SF to examine the acceptability of five behavioral treatments (positive reinforcement, chair time-out, response cost, overcorrection, and reprimands) for general or medically related noncompliance of a hypothetical child described as either healthy or with cancer. Ratings were obtained from parents of children receiving medical treatment for cancer, pediatric nurses, and parents of medically healthy children. Results revealed no significant differences in acceptability ratings based on the type of noncompliance (general vs. medically related) or based on the child being described as healthy or with cancer. Parents of children receiving medical treatment for cancer rated the response cost and the time-out procedure significantly lower than the participants. Parents of healthy children rated the positive reinforcement procedure significantly less acceptable than the other groups.

Fairbanks and Stinnett (1997) evaluated treatment acceptability of different behavior interventions associated with different diagnostic labels as rated by members of different professional groups. A vignette was presented to 31 teachers, 33 school psychologists, and 33 school social workers from three school districts in southwestern Illinois. A copy of the IRP-15 was completed by the participants to rate the acceptability of the proposed treatment described within the vignette. Variables manipulated included the diagnostic label of the child described within the case vignette. Diagnostic labels manipulated were learning disabled (LD), behavior disordered (BD), and Attention Deficit Disorder (ADD). The case vignette described a third grade boy who displayed behaviors such as excessive talking, out-of-seat, and overly active. A three-way analysis of variance was used to evaluate the data. Findings indicated no significant differences based on the diagnostic label of the child described in the case vignette.

In summary, research has shown that client variable manipulations may influence results concerning treatment acceptability. At the same time, there has been research that has illustrated that consumer variable manipulation can affect perceptions of treatment acceptability.

Rater Variable Manipulation

A third line of research identified within the literature on treatment acceptability involved examining the impact of different characteristics of raters. Prior to 1990, examination of different characteristics of raters involved: (a) gender of raters (Kazdin, 1980), (b) raters’ knowledge of treatment (Singh & Katz, 1985), and (c) affiliation of raters (Heffer & Kelley, 1987). Recent variable examination of different raters has involved: (a) geographic location of raters’ high schools (Stinnett et al., 2001), (b) raters’ knowledge of treatment (Gage & Wilson, 2000; Rasnake, Martin, Tarnowski, & Mulick, 1993; Singh & Katz), and (c) professional affiliation of raters (Fairbanks & Stinnett, 1997; Spreat & Walsh, 1994).

Gender of raters. Supplementary analyses within Kazdin’s (1980) study focused on the gender of the raters and revealed a statistically significant difference in male ratings versus female ratings, with males rating electric shock as more acceptable than females and reinforcement as less acceptable than females. The hierarchical ordering of treatments did not differ by gender. Miller and Kelley (1992) evaluated three rater variables (gender, marital adjustment, and child behavior) on treatment acceptability. The TEI was administered to 69 married couples with young children to assess the acceptability of six interventions (positive reinforcement, response cost, medication, room timeout, chair timeout, and spanking) for a hypothetical child displaying noncompliance and aggressiveness. The couples also completed a marital adjustment scale and a child behavior problem inventory. Findings revealed that gender influenced treatment acceptability ratings for all interventions except chair time-out, with mothers providing higher ratings than fathers for all interventions except spanking and medication. When compared to non-distressed couples, those couples in marital distress provided significantly higher acceptability ratings for room time-out and significantly lower ratings for positive reinforcement. Couples of children
displaying problem behaviors rated medication as more acceptable and spanking as less acceptable than couples of children without behavior problems.

Raters’ knowledge of treatment. Singh and Katz (1985) presented case descriptions to 96 undergraduate psychology students and asked them to rate the acceptability of four different treatments (differential reinforcement of incompatible behavior, positive practice overcorrection, time-out, and humanistic parenting) using the TEI. Results indicated the following ranking for acceptability with the greatest acceptability first: differential reinforcement of incompatible behavior, humanistic parenting, and positive practice. Time-out received ratings as least acceptable. Following this study, the same participants were provided formal educational training on three behavioral treatments (differential reinforcement of incompatible behavior, positive practice overcorrection, and time-out) and then the participants re-evaluated all four treatments. The training provided information such as specific details of each treatment, empirical data outlining the effectiveness of each treatment, and potential side effects of the treatments. At post-training, the participants’ treatment rating revealed the rankings to be differential reinforcement of incompatible behavior, positive practice overcorrection, and time out. The humanistic parenting was rated lower at post education than at pre-education.

Rasnake et al. (1993) examined the influence of knowledge of behavioral principles on treatment acceptability. They administered the IRP to 57 institutional staff members to evaluate the acceptability of six treatments (differential reinforcement of other behavior, differential reinforcement of incompatible behavior, stimulus control, overcorrection, physical restraint, and contingent shock) for self-injurious behavior. In addition, the staff members completed a measure of their understanding of general behavioral principles. The stimulus control procedure was rated as the most acceptable of the treatments while the other treatments were not rated as acceptable treatments. Findings revealed that older residential staff rated the stimulus control procedure as significantly more acceptable than younger residential staff. Neither variables of knowledge of general behavioral principles or work experience were found to influence acceptability ratings.

Affiliation of raters. Heffer and Kelley (1987) found that specific affiliations of raters influenced ratings of treatment acceptability. They used the TEI to assess mothers’ ratings of five child management techniques. They found that raters from different socioeconomic classes and from different races rated acceptability of treatments differently. In a previously described study, Spreat and Walsh (1994) found differences in treatment acceptability among members of different divisions of the American Association on Mental Retardation (AAMR) assessed factors associated with decisions regarding acceptability of behavior modification programs by members from differing divisions of the AAMR. Findings indicated that the strongest indicator of treatment acceptability was the respondents’ personal estimates of probable treatment success. Members of the Psychology Division of the AAMR rated treatments as slightly more acceptable than members of other divisions of AAMR.

In a previously described study by Fairbanks and Stinnett (1997), participants evaluated treatment acceptability of different behavior interventions as rated by members of three different professional groups. Findings indicated that teachers rated the negative intervention (time-out from reinforcement with praise) more acceptable than school psychologists or school social workers. Also, school psychologists rated the negative intervention more acceptable than school social workers.

Waas and Anderson (1991) used the CIRP to reveal differences in treatment acceptability ratings of 2nd graders, 5th graders, and college students. Participants rated the acceptability of a behavior contingency intervention, group counseling, and special education class placement. Findings revealed that college students rated the behavior contingency intervention and special education class placement significantly lower than the 2nd and 5th graders. No significant differences in acceptability ratings were found for the group counseling treatment.

Geographic location of raters’ high school. In a previously described study in which diagnostic label of the client was also examined, Stinnett
et al. (2001) evaluated the impact of geographic location of raters’ high schools on treatment acceptability. Teachers-in-training were asked information regarding the high school from which they graduated and grouped by graduation from a rural versus an urban high school. Findings indicated that location of the high school of raters influenced acceptability of treatments, with participants who graduated from rural high schools rating treatments as more acceptable than participants who graduated from urban high schools.

### TABLE 2
Summary of Recent (1990–2005) Treatment Acceptability Research

<table>
<thead>
<tr>
<th>Studies</th>
<th>Treatment Variable Manipulation</th>
<th>Client Variable Manipulation</th>
<th>Rater Variable Manipulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tingstrom, 1990</td>
<td>Mild vs. severe problem behavior revealed significantly higher acceptability of a time-out procedure for more severe problem behavior</td>
<td>Teacher vs. school psychologist ratings of time-out revealed no significant differences on acceptability</td>
<td>2nd graders, 5th graders, and college student revealed similarities and differences</td>
</tr>
<tr>
<td>Waas &amp; Anderson, 1991</td>
<td>2nd graders, 5th graders, and college student revealed similarities and differences</td>
<td>2nd graders, 5th graders, and college student revealed similarities and differences</td>
<td></td>
</tr>
<tr>
<td>Miller &amp; Kelley, 1992</td>
<td>Behavior interventions and medication revealed similarities and differences</td>
<td>Mothers vs. fathers, maritally distressed couples vs. non-maritally distressed couples, and parents of children with and without behavior problems revealed similarities and differences</td>
<td></td>
</tr>
<tr>
<td>Reimers et al., 1992</td>
<td>Mildly intrusive treatments rated more acceptable for mild problem behaviors and highly intrusive treatments rated more acceptable for severe problem behaviors</td>
<td>Analog case descriptions and clinical cases revealed similar ratings</td>
<td></td>
</tr>
<tr>
<td>Betts &amp; Remer, 1993</td>
<td>Paradoxical treatment found less acceptable than nonparadoxical treatment</td>
<td>Age of institutional staff influenced acceptability; knowledge of behavioral principles and years of work experience did not</td>
<td></td>
</tr>
<tr>
<td>Rasnake et al., 1993</td>
<td>Differences and similarities revealed among six treatments for self-injurious behavior</td>
<td>Age of institutional staff influenced acceptability; knowledge of behavioral principles and years of work experience did not</td>
<td></td>
</tr>
</tbody>
</table>
Gage and Wilson (2000) randomly assigned 30 parents of children with ADHD (as determined by the Child Behavior Check List and a diagnosis by a professional) and 30 parents of children without an ADHD diagnosis to three different conditions. The three experimental conditions consisted of: (a) medication, (b) behavioral treatment, and (c) a combination...
of both. Following a case vignette, the TEI was used to measure acceptability of treatments. Data were analyzed using three ANOVAs for each treatment scenario across TEI ratings. Results indicated that parents of children with ADHD rated medications and combination of both treatments statistically significantly higher than parents of children without ADHD. Parents of children without ADHD rated the behavioral treatment statistically significantly higher than parents of children with ADHD. Parents of children with ADHD rated the combination of treatments statistically significantly higher than the other treatments. No other significant effects were found within the parents of children with ADHD group. Among parents of children without ADHD, results revealed statistically significant differences among all treatments with the behavioral treatment receiving the highest acceptability treatment, followed by the combination treatment, and then the medication treatment.

Summary of Research on Treatment Acceptability

A summary of the research on treatment acceptability reveals that different types of treatment may produce different levels of treatment acceptability (Betts & Remer, 1993; Kazdin, 1980). The examination of the individual implementing a treatment (Kalfus & Burk, 1989) appears to have been replicated (Tingstrom, 1990) with consistent findings. The setting in which a treatment is proposed to be implemented has also been found to influence acceptability ratings (Burgio et al., 1995). Several other factors that have been found to impact the acceptability of some treatments include level of restrictiveness of treatments (Kazdin; Spreat & Walsh, 1994), information on positive outcome of treatments (Clark & Elliott, 1988), treatment history/severity (Kalfus & Burk; Kazdin; Spreat & Walsh), location of rater’s high school (rural vs. urban; Stinnett et al., 2001), raters’ professional affiliation (Fairbanks & Stinnett, 1997; Spreat & Walsh), professional recommending treatment (Carter, 2005), and raters’ experiences as a parent (having or not having a child diagnosed with ADHD; Gage & Wilson, 2000). Factors that have been found to have no effect on treatment acceptability include treatment mediators (Kalfus & Burk), client age/gender (Elliott & Fuqua, 2002), client labeling (Fairbanks & Stinnett; Stinnett et al.), and increasing raters’ knowledge of treatment (Rasnake et al., 1993). Refer to Table 2 for recent studies reviewed.

With the availability of numerous treatments for problem behaviors treatment variables appear to be the most frequently examined variables within the literature on treatment acceptability although directly examining acceptability of different treatments in isolation from other variable manipulations rarely appears as the primary focus of recent research. Alberto and Troutman (1999) have stated that a hierarchy of intrusiveness among behavior interventions is commonly accepted within the literature. Treatments that are considered more intrusive are generally considered less acceptable, but can be influenced by descriptions of client problem severity (Miltenberger, 1990).

Client/case description variables that appear to consistently influence acceptability ratings include the severity of the problem behavior being described. While in general, higher acceptability ratings are typical for more severe problem behavior, the research has shown a tendency for higher acceptability ratings when more intrusive treatments are matched to more severe behaviors and less intrusive treatments to less severe behaviors. Additionally, recent literature has evaluated several client/case description variables that do not appear to consistently influence acceptability ratings. These variables include the sex of the client, age of the client, diagnostic label of the client, and descriptions of client health.

Rater variables have been extensively examined in the recent literature and numerous interactions have been shown between rater variables and the acceptability of specific treatments. Some of the rater variables found to influence acceptability ratings include gender, age, current educational level, membership in a specific division of a professional organization, personal estimates of treatment success, marital distress, parenting a child with behavior problems, professional title (teacher vs. school psychologist), parenting a child with or without a health problem, graduating from a rural or urban high school. Knowledge of
behavioral principles and years of work experience were not found to influence acceptability ratings.

It appears that researchers have considered numerous factors that could influence the acceptability of treatments. The manipulation of variables related to treatments, clients, raters has shown that all types of variable manipulations can influence the acceptability of ratings of treatments. With the advances that have been made with treatment acceptability instruments, the evaluation of specific treatments, specific populations of clients, and raters from specific areas of the country with varying backgrounds should become more prevalent. Miltenberger (1990) recommended that professionals could collect representative acceptability ratings from the area or school in which they work for a number of different treatments and problem behaviors. The professional could then use this information toward making treatment decisions. With the improved instruments for assessing treatment acceptability and the advances in treatment acceptability research, professionals have a well-developed basis from which to make treatment decisions and conduct further research.

References


Received: 6 March 2006
Initial Acceptance: 20 April 2006
Final Acceptance: 15 July 2006