Research on professionals', consumers', and others' judgments of the restrictiveness, social acceptability, and estimated frequency of use of procedures to decrease behavior was reviewed. General findings were that (a) respondents generally were consistent in rating procedures from least to most restrictive; (b) most respondents agreed that procedures judged more restrictive should be used as a last resort; (c) more restrictive procedures were not frequently used in practice; (d) respondents generally agreed that less restrictive procedures were more socially acceptable, and vice versa; and (e) the social acceptability of procedures changed as a function of contextual variables. Problems related to research methodology were discussed and future research directions suggested.

The use of procedures to decrease behavior (i.e., aversive procedures) has generated considerable controversy. Rights advocates have criticized the use of such procedures as unjustifiable for ethical reasons (Guess, Helmstetter, Turnbull, & Knowlton, 1987; Turnbull et al., 1986). Legal and behavioral professionals have admonished those who use the procedures for their apparent insensitivity to issues of restrictiveness and social acceptability (Friedman, 1975; Guess et al., 1987; LaVigna & Donnellan, 1986). Restrictiveness has been defined as judgments of the extent to which a procedure curtails one's freedom (Budd & Baer, 1976). Social acceptability has been defined as judgments by lay persons, clients, and others of whether behavioral procedures are appropriate, fair, and reasonable for the problem or client (Kazdin, 1981). In contrast, those who recommend controlled use of the procedures have pointed to empirical research demonstrating their effectiveness in decreasing severely maladaptive behaviors (Axelrod & Apsche, 1983; Foxx, 1982; Matson & Taras, 1989; Van Houten et al., 1988).

The controversy over procedures to decrease behavior has provided the impetus to examine research related to judgments of the restrictiveness, social acceptability, and frequency of use of these procedures in treatment settings. These
Criteria Used in Evaluating Procedures to Decrease Behavior

LaVigna and Donnellan (1986) recommended that multiple criteria be used in evaluating procedures employed to decrease behavior, including effectiveness, restrictiveness, and social acceptability. The effectiveness of procedures to decrease behavior has been examined elsewhere (e.g., Axelrod & Apsche, 1983; Guess et al., 1987). Several contextual variables must be considered in examining the literature on restrictiveness and social acceptability, including (a) circumstances within which the behavior occurred, (b) severity of the behavior, (c) a person's expertise in administering the procedures, (d) reported effectiveness of the behavioral procedure, (e) adverse side effects. These contextual variables may have dramatic effects on judgments of a procedure's restrictiveness or social acceptability. For example, using a time-out procedure in a seclusion room as a consequence for a child's noncompliant behavior might be judged highly restrictive and socially unacceptable in this context. Yet, the same procedure might be judged less restrictive and more socially acceptable in the context of physically aggressive behavior that threatened the safety of the child or others.

Research Review on Judgments of Behavioral Procedures

Studies were identified from a computer search of the Current Index of Journals in Education and Psychological Abstracts for the years 1975 through 1987. Descriptors included aversion conditioning, aversive stimulation, aversion therapy, behavior management, behavior modification, developmental disabilities, legislation, legal responsibility, mental retardation, punishment, and responsibility. Additional studies were identified by searching the references listed in the primary studies and by following up on recommendations made by a journal reviewer. Also, several studies on the social acceptability of behavioral procedures reviewed by Reimers, Wacker, and Koeppl (1987) were included.

The literature on restrictiveness and social acceptability of behavioral procedures was divided into four categories: (a) opinion articles by legal or behavioral professionals, who arranged procedures hierarchically according to restrictiveness; (b) survey studies of professionals working with individuals having developmental disabilities, which yielded data on judgments of procedures' restrictiveness; (c) analogue studies of judgments by professionals, consumers, and others, which yielded ratings of procedures' social acceptability; and (d) survey studies of professionals on the judged frequency of procedures' use.

Hierarchies of Procedures Based on Legal or Behavioral Opinion

Legal and behavioral professionals (Barton, Brulle, & Repp, 1983; Friedman, 1975) have advocated the development of hierarchies of behavioral procedures based on the concept of least restriction. This concept is based on the philosophy and legal precedent that procedures should be ranked according to their restrictiveness and that procedures judged more restrictive should only be used when less restrictive ones are shown to be ineffective (Foxx, 1982; Friedman, 1975; Wyatt v. Stickney, 1972). Table 1 presents hierarchies that were arranged by their authors, without empirical data from other professionals or consumers, in levels from least to most restrictive. Most hierarchies only included procedures used to decrease behavior, although some also included procedures to increase behavior (i.e., positive reinforcement). The procedures were defined and/or described in most hierarchies (Bimbrauer, 1978; Brakman, 1985; Foxx, 1982;
Table 1
Comparison of Hierarchies Based on Legal or Behavioral Opinion

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Modeling, redirection, positive reinforcement, token economy, timeout (within view); extinction, response cost, graduated guidance</td>
<td>1. Differential reinforcement of other behavior, extinction</td>
<td>1. Extinction</td>
<td>1. Differential reinforcement of other behavior, differential reinforcement of incompatible behavior, satiation</td>
<td>1. Differential reinforcement of other behavior</td>
</tr>
<tr>
<td>2. Timeout (out of view), response cost (loss of activity), positive practice, restitutional overcorrection</td>
<td>2. Extinction</td>
<td>2. Verbal reprimand</td>
<td>2. Extinction</td>
<td></td>
</tr>
<tr>
<td>3. Required relaxation, food removal, timeout (out of room)</td>
<td>3. Response cost, overcorrection, restitution</td>
<td>3. Response cost</td>
<td>3. Verbal aversive (e.g., &quot;No!&quot;)</td>
<td></td>
</tr>
<tr>
<td>5. Electric shock, food/water deprivation, physical striking, environmental extremes</td>
<td>5. Classroom timeout</td>
<td>5. Classroom timeout</td>
<td>5. Timeout</td>
<td></td>
</tr>
</tbody>
</table>

Note: Hierarchies are listed in order of increasing restrictiveness.

Nelson & Rutherford, 1983). All authors included the contextual variable of accounting for the severity of the maladaptive behavior. No hierarchies included exactly the same set of behavioral procedures. Generally, procedures common to hierarchies were positioned at similar levels of restrictiveness. For example, positive reinforcement (or differential reinforcement) procedures were included and ranked least restrictive in the hierarchies by Birnbrauer, Foxx, Heads, and Nelson and Rutherford. Conversely, electric shock and physical striking were included and ranked among the most restrictive by Birnbrauer and Heads. Procedures such as extinction, response cost, and various forms of time-out were frequently ranked among the intermediate levels. Although there was considerable commonality across hierarchies, some procedures were ranked at different levels. For example, overcorrection was ranked at least restrictive levels by Brakman and Heads, but at more restrictive levels by Birnbrauer, Nelson and Rutherford, and Foxx.

Hierarchies of behavioral procedures based on professional opinion must be evaluated cautiously. One potential problem might be that practical differences in restrictiveness can be related to the various ways that a procedure is implemented or the severity of the maladaptive behavior to which it is applied. Thus, the generic term used to identify a behavioral procedure (e.g., time-out, overcorrection) may not provide substantive information about its restrictiveness in actual implementation. These contextual issues were mentioned by some authors as limitations of the hierarchies (e.g., Heads, 1978; Nelson & Rutherford, 1983). Perhaps the most significant limitation of these hierarchies, however, was that they were not based on empirical data of professionals’ or consumers’ judgments.

Hierarchies of Procedures Based on Survey Studies

In three surveys, professional respondents judged the intrusiveness or restrictiveness of procedures used to decrease behavior (Augustine & Cipani, 1982; Irvin & Lundervold, 1988; Morgan & Striefel, 1987–1988). The authors interpreted results by arranging the procedures in hierarchies on the evaluated dimension.

Augustine and Cipani (1983) distributed a survey questionnaire to nine members of a behavioral management committee at a center serving individuals with developmental disabilities to determine the intrusiveness of 13 behavioral procedures used to decrease self-injurious behavior. Intrusiveness was not defined by the investigators but was defined elsewhere as the extent to which an intervention involves either physical or

Morgan
psychological unpleasantness (Irvin & Lundervold, 1988). The concept of intrusiveness has been considered similar to the concept of restrictiveness (Friedman, 1975; Irvin & Lundervold, 1988). Augustine and Cipani described procedures and specified the length of time that each procedure was to be applied. A description of their study is presented in Table 2.

Because Augustine and Cipani did not define intrusiveness, their respondents may have rated procedures on different, subjectively determined dimensions. Also, respondents were apparently not provided with information on the effectiveness of procedures, potentially adverse side effects, availability of positive reinforcement for incompatible behaviors, the expertise of the individual administering the procedure, or other contextual information.

Morgan and Striefel (1987–1988) examined the judgments of restrictiveness of 13 procedures to decrease behavior. Respondents were told that the procedures would be used in conjunction with positive reinforcement for incompatible in a treatment program. Definitions and examples of the procedures were provided. However, the definition of restrictiveness, a description of the maladaptive behavior to be decreased, and other information were not included. The ranking

Table 2
Descriptions and Results of Studies Examining Intrusiveness and Restrictiveness

<table>
<thead>
<tr>
<th>Study</th>
<th>No. and type of respondent/No. of procedures</th>
<th>Information given to respondents</th>
<th>Dimension(s) used for evaluation</th>
<th>Ordering of procedures*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Augustine &amp; Cipani</td>
<td>9 psychologists, direct-care staff, social worker, nurse, administrator/13</td>
<td>Case was described as an institution resident with self-injurious behavior and severe mental retardation</td>
<td>Rating of intrusiveness</td>
<td>Extinction, nonexclusionary timeout (TO), response cost (RC), token RC, brief immobilization, positive practice overcorrection (PPO), lemon juice/citric acid, water mist, icing, seclusion TO, aromatic ammonia, physical restraint, and electric shock</td>
</tr>
<tr>
<td>(1982), Experiment 3</td>
<td></td>
<td></td>
<td></td>
<td>Changing antecedent events, differential reinforcement of incompatible behavior (DRI), planned ignoring, RC, nonexclusionary TO, PPO and restitutional overcorrection (RO) without physical guidance, PPO and RO with physical guidance, exclusionary TO, contingent physical restraint, discomforting stimuli, and electric shock</td>
</tr>
<tr>
<td>Morgan &amp; Striefel</td>
<td>118 psychologists, administrators, teachers, specialists/13</td>
<td>Positive reinforcement was concurrently available while a procedure to decrease behavior was implemented</td>
<td>Ranking of restrictiveness</td>
<td></td>
</tr>
<tr>
<td>(1987–1988)</td>
<td></td>
<td></td>
<td></td>
<td>(Partial lists of procedures rated): Restrictiveness: differential reinforcement (DR), simple correction, social disapproval, physical and mechanical restraint, seclusion TO; Intrusiveness: DR, simple correction, extinction . . . contingent noxious stimulation (CNS), mechanical restraint, electric shock; and Acceptability: electric shock, mechanical restraint, CNS . . . extinction, simple correction, DR</td>
</tr>
<tr>
<td>Irvin &amp; Lundervold</td>
<td>58 special education teachers, program coordinators, administrators/18</td>
<td>Definitions of each of four dimensions</td>
<td>Ratings of restrictiveness, intrusiveness, acceptability, and efficacy</td>
<td></td>
</tr>
<tr>
<td>(1988)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Least to most.
method differed from the bipolar rating method employed in other studies and allowed the examination of agreement among rankers on the placement of procedures. A hierarchy of behavioral procedures was generated by the authors using frequencies of ranks as the dependent measure (see Table 2). General agreement was found in the rankings of some procedures. For example, most respondents agreed that electric shock was most restrictive. However, less agreement was found for some other procedures, such as exclusionary time-out and positive practice overcorrection with physical guidance.

Irvin and Lundervold (1988) conducted a survey with 58 special educators. These investigators examined the educators' ratings of restrictiveness, intrusiveness, acceptability, and efficacy of behavioral procedures to decrease behavior. Each of the four dimensions was defined for the respondents. No contextual information was provided. Results indicated that consensually high and low ratings were obtained for acceptability, restrictiveness, and intrusiveness. That is, highly acceptable procedures were rated low on restrictiveness and intrusiveness, and vice versa. As shown in Table 2, differential reinforcement, simple correction, and social disapproval were consistently rated most acceptable, least restrictive, and least intrusive. Conversely, contingent electric shock, mechanical restraint, and contingent noxious stimulation were consistently rated least acceptable, most restrictive, and most intrusive. Seclusion and exclusionary time-out were also among those procedures rated most restrictive and most intrusive.

In the Irvin and Lundervold (1988) study, ratings of restrictiveness and intrusiveness were positively correlated. In fact, based on the correlation data, the investigators concluded that the dimensions of intrusiveness and restrictiveness had not been differentiated by the respondents, despite different definitions that emphasized degree of discomfort (i.e., intrusiveness) versus extent of curtailed access (i.e., restrictiveness).

Although there was general correspondence in the hierarchical placements of behavioral procedures common to the three studies just described, contextual variables that may have influenced judgments were not considered. For example, none of the investigators examined the potential influences of severity of behavior, expertise of the person administering the procedure, or adverse side effects. However, some of these variables were addressed by researchers on the social acceptability judgments of behavioral procedures.

Research on the Social Acceptability of Behavioral Procedures

Judgments by treatment consumers and others to determine the social acceptability of behavioral procedures have been examined in several studies. Summaries of 20 investigations are presented in Table 3. In these analogue studies, the respondents were asked to review a case description involving an individual whose maladaptive behavior was to be decreased. The case was usually presented in written form, unless otherwise indicated in Table 3. The respondent then read a description of one or more behavioral procedures and rated social acceptability. An average social acceptability rating score was computed for each procedure. Although hierarchies of procedures' social acceptability were not described, some investigators analyzed results by arranging a rank ordering of the procedures.

Respondents included special education teachers in four studies (Elliott, Witt, Galvin, & Peterson, 1984; Martens, Witt, Elliott, & Darveau, 1985; Von Brock & Elliott, 1987; Witt, Martens, & Elliott, 1984), institution and community-based program staff in one study (Miltenberger, Lennox, & Erfanian, 1989), and parents of children receiving treatment in five studies (Kazdin, 1984; Kazdin, French, & Sherick, 1981; Miltenberger, Parrish, Rickert, & Kohr, in press); Pickering & Morgan, 1985; Singh, Watson, & Winton, 1987). Respondents in other studies included regular education teachers, student and preservice teachers, parents of children who were not in treatment, children in inpatient psychiatric treatment, children who were not in treatment, and undergraduate college students (see Table 3).

Cases described in each study were individuals with maladaptive behaviors to be decreased (e.g., and 8-year-old girl with aggressive and oppositional behavior). In eight studies, the cases were described as individuals with mental retardation (Kazdin, 1980b, 1981, 1984; Kazdin et al., 1981; Miltenberger et al., 1989; Pickering & Morgan, 1985; Singh & Katz, 1985; Singh et al., 1987). In most studies, respondents read a written description of one procedure used to decrease the individual's behavior. Then, they rated the procedure's social acceptability using an assessment instrument, such as the Treatment Evaluation Inventory (Kazdin, 1980a), Semantic Differential
<table>
<thead>
<tr>
<th>Study</th>
<th>Types/No. of respondents</th>
<th>Procedures evaluated</th>
<th>Case descriptions</th>
<th>Assessment instrument used</th>
<th>Summary of findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kazdin (1980a), two experiments</td>
<td>Undergraduate college students/144 in each experiment</td>
<td>Differential reinforcement of incompatible behavior (DRI), nonexclusionary timeout (TO), exclusionary TO</td>
<td>Case was described as either an 8- or 10-year-old child with disruptive behavior</td>
<td>Treatment Evaluation Inventory (TEI), Semantic Differential (SD)</td>
<td>Experiment 1: DRI and nonexclusionary TO were more acceptable than isolation; Experiment 2: isolation was more acceptable when used as a “backup” procedure and with a contingency contract.</td>
</tr>
<tr>
<td>Kazdin (1980b), two experiments</td>
<td>Undergraduate college students/88 in each experiment</td>
<td>DRI, TO, medication, electric shock</td>
<td>Case was described as either 5- or 10-year-old child with behavior problems such as opposition, aggression, self-injury, etc.; some children had mental retardation</td>
<td>TEI, SD</td>
<td>DRI was most acceptable in both experiments, followed by TO, medication, and electric shock; case descriptions of severe behavior problems or mental retardation did not alter the rated order of treatments’ acceptability.</td>
</tr>
<tr>
<td>Kazdin (1981), two experiments</td>
<td>Undergraduate college students/112 in each experiment</td>
<td>DRI, positive practice overcorrection (PPO), TO, medication</td>
<td>Case was described as either 8- or 10-year-old child with behavior problems; one case description involved child with mental retardation</td>
<td>TEI, SD</td>
<td>DRI was most acceptable in both experiments, followed by PPO, TO, and medication; effectiveness did not influence acceptability ratings, but adverse side effects reduced acceptability.</td>
</tr>
<tr>
<td>Kazdin et al. (1981)</td>
<td>Child psychiatric patients/32, parents/32, psychiatric staff/32</td>
<td>DRI, PPO, medication, TO</td>
<td>(Same as Kazdin, 1981 above)</td>
<td>TEI, SD</td>
<td>DRI was most acceptable. TO was least acceptable; children rated treatments as less acceptable than did parents.</td>
</tr>
<tr>
<td>Witt &amp; Martens (1983)</td>
<td>Student teachers/180</td>
<td>Praise, ignoring, home-based reinforcement, response cost (RC), seclusion, TO, and token economy</td>
<td>Fifth-grade boy with either mild, moderate, or severe behavior problem</td>
<td>Intervention Rating Profile (IRP)</td>
<td>Positive procedures applied to mild behavior problems requiring low amounts of teacher time were more acceptable; procedures’ acceptability depended on severity of problem.</td>
</tr>
<tr>
<td>Elliott, Witt, Galvin, &amp; Peterson (1984)</td>
<td>Regular and special education teachers/70 and 71 in Experiments 1 and 2, respectively</td>
<td>Praise, ignoring, home-based reinforcement, RC, seclusion</td>
<td>Behavior problem described as either mild (i.e., daydreaming), moderate (i.e., obscene language), or severe (i.e., property destruction)</td>
<td>IRP</td>
<td>(a) Acceptability ratings of both positive and reductive procedures varied with severity of the problem behavior, (b) less complex or time-consuming procedures were rated more acceptable, (c) positive procedures were more acceptable than negative procedures.</td>
</tr>
<tr>
<td>Kazdin (1984)</td>
<td>Children ages 7 to 12 in inpatient therapy/40; parents/40</td>
<td>TO in classroom or one’s own room, locked seclusion, medication</td>
<td>Case was described as either 8-year-old girl with aggressive and oppositional behavior, or 10-year-old boy with mental retardation and hyperactive and disruptive behavior</td>
<td>TEI (child and adult versions), SD</td>
<td>Children’s order of acceptability: (1) medication, (2) TO, (3) seclusion; parent’s order of acceptability: (1) TO, (2) seclusion, (3) medication; procedures described as more effective were rated as more acceptable by both children and parents.</td>
</tr>
<tr>
<td>Study</td>
<td>Types/No. of respondents</td>
<td>Procedures evaluated</td>
<td>Case descriptions</td>
<td>Assessment instrument used</td>
<td>Summary of findings</td>
</tr>
<tr>
<td>-------</td>
<td>-------------------------</td>
<td>----------------------</td>
<td>-------------------</td>
<td>---------------------------</td>
<td>---------------------</td>
</tr>
<tr>
<td>Witt, Martens, &amp; Elliott (1984)</td>
<td>Regular and special education teachers from grades K through 12 with 1 to 25 yrs. experience/180</td>
<td>Praise, ignoring, home-based reinforcement, RC, token economy, seclusion TO</td>
<td>Fifth-grade boy with either mild (daydreaming), moderate (obscene language), or severe (property destruction) behavior problem</td>
<td>IRP</td>
<td>Teacher's judgments of acceptability were influenced by the time required to implement the procedure, e.g., more acceptable procedures were those that required little time and were applied to mild behavior problems (a) Positive procedures were a more acceptable use of teacher time than negative procedures, (b) severity of behavior problem did not influence ratings of acceptability (a) Procedures were more acceptable when applied to severe behavior problem, (b) no statistically significant difference between written and videotaped case description</td>
</tr>
<tr>
<td>Witt, Elliott, &amp; Martens (1984)</td>
<td>Preservice and student teachers/180</td>
<td>(Same as Witt et al., 1984, above)</td>
<td>(Same as Witt et al., 1984, above)</td>
<td>IRP</td>
<td></td>
</tr>
<tr>
<td>Martens, Witt, Elliott, &amp; Darveau (1985)</td>
<td>Regular and special education teachers/54</td>
<td>RC, RC plus going to principal's office</td>
<td>Fifth-grade boy with either mild (daydreaming) or severe (property destruction) behavior problem. Case presented via written description or videotape</td>
<td>IRP-15 (i.e., shortened version of IRP) and SD</td>
<td></td>
</tr>
<tr>
<td>Pickering &amp; Morgan (1985)</td>
<td>Parents of children with autism: 13; parents of children with other handicaps: 33; parents of children without handicaps: 73</td>
<td>DRI, TO, overcorrection (OC), electric shock</td>
<td>(Same as Kazdin, 1980b, above)</td>
<td>TEI, SD</td>
<td>(a) All parent groups rated DRI most acceptable, (b) shock was rated most likely to result in discomfort and negative side effects, (c) overcorrection was rated less acceptable by parents of children without handicaps (a) DRI was initially more acceptable than others, (b) after education and instruction in the use of behavioral procedures, all were rated more acceptable</td>
</tr>
<tr>
<td>Singh &amp; Katz (1985)</td>
<td>College students/96</td>
<td>DRI, PPI, and TO; humanistic parenting was included as a control procedure</td>
<td>Case described was either 16-year-old male with mental retardation and hyperactivity or 8-year-old girl with aggressive behaviors</td>
<td>TEI, SD</td>
<td></td>
</tr>
<tr>
<td>Witt &amp; Robbins (1985), two experiments</td>
<td>Experiment 1: regular education, preschool, and Headstart teachers/96; Experiment 2: regular education teachers/98</td>
<td>Differential reinforcement of other behavior (DRO) and of low rates (DRL), TO, corporal punishment, reprimands, staying after school</td>
<td>(Same as Witt et al., 1984, above)</td>
<td>IRP</td>
<td>Experiment 1: DRO was most acceptable, corporal punishment was least acceptable; Experiment 2: procedure implemented by teacher was more acceptable than procedure implemented by principal; less experienced teachers rated procedures more acceptable than more experienced teachers</td>
</tr>
<tr>
<td>Elliott, Witt, Galvin, &amp; Moe (1986), two experiments</td>
<td>Sixth graders without handicaps/23 and 79 in Experiments 1 and 2, respectively</td>
<td>Positive and negative individual and group contingencies delivered verbally, through loss/gain of recess time, or as &quot;traditional&quot; consequences</td>
<td>A peer in the classroom who talked out of turn or destroyed others' property</td>
<td>Experiment 1: questionnaire that assessed severity of behavior problems and ratings of procedures; Experiment 2: IRP adapted for children</td>
<td>Experiment 1: positive individual and group contingencies and negative individual contingencies were most acceptable, Experiment 2: positive group and negative individual contingencies were most acceptable</td>
</tr>
<tr>
<td>Study</td>
<td>Types/No. of respondents</td>
<td>Procedures evaluated</td>
<td>Case descriptions</td>
<td>Assessment instrument used</td>
<td>Summary of findings</td>
</tr>
<tr>
<td>------------------------------</td>
<td>----------------------------------------------------------------</td>
<td>---------------------------------------</td>
<td>----------------------------------------------------------------------------------</td>
<td>-----------------------------</td>
<td>-------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Frentz &amp; Kelley (1986)</td>
<td>Mothers of children ages 2 to 12 not receiving special services/82</td>
<td>Differential attention, RC, spanking, TO, TO with spanking</td>
<td>Eight-year-old boy with noncompliant or aggressive behavior problems described as either mild or severe</td>
<td>TEI, Eyberg Child Behavior Inventory</td>
<td>(a) RC was most acceptable; (b) TO was more acceptable than differential attention, TO with spanking, and TO with spanking alone; (c) all procedures were rated more acceptable when used with severe behavior problems</td>
</tr>
<tr>
<td>Martens, Peterson, Witt, &amp; Cirone (1986)</td>
<td>Regular and special education teachers/2,279</td>
<td>Positive and negative classroom procedures were described in a 65-item intervention profile</td>
<td>Not applicable</td>
<td>Questionnaire, i.e., Classroom Intervention Profile</td>
<td>(a) Teachers' responses factored into categories; (b) teachers differentially rated categories in terms of effectiveness, ease of use, and frequency of use of procedures; (c) procedures involving redirection or use of rewards were most acceptable</td>
</tr>
<tr>
<td>Singh, Watson, &amp; Winton (1987)</td>
<td>Mothers of children with mental retardation/96</td>
<td>DRI, OC, TO, drug therapy</td>
<td>Case described was 16-year-old male with hyperactivity or 6-year-old girl with aggressive behavior. Both had mental retardation</td>
<td>TEI, SD</td>
<td>DRI was most acceptable, followed by OC; TO and drug therapy were least acceptable</td>
</tr>
<tr>
<td>Von Brock &amp; Elliott (1987)</td>
<td>Regular and special education teachers/216</td>
<td>Token economy, TO, RC</td>
<td>Fifth-grade boy with either mild (daydreaming) or severe (property destruction) behavior problem</td>
<td>Behavior Intervention Rating Scale (BIRS), SD</td>
<td>(a) Token economy and RC were more acceptable than TO; (b) information on effectiveness influenced ratings of procedures' acceptability for mild problems, but not severe; (c) less acceptable procedures were judged less effective</td>
</tr>
<tr>
<td>Mittenberger, Lennox, &amp; Efranian (1989), two experiments</td>
<td>Experiment 1: community residential facility staff/72; Experiment 2: institution staff/40</td>
<td>Both experiments: DRO, TO, OC, electric shock</td>
<td>Both experiments: young adult residents with severe mental retardation and either mild (i.e., tantrums) or severe (i.e., aggressive) behaviors</td>
<td>TEI</td>
<td>Experiment 1: (a) DRO was most acceptable, followed by TO, OC, and shock. (b) DRO was rated more acceptable for mild than severe problems. Experiment 2: (a) same order of acceptability, except no difference between DRO and TO; (b) shock was rated more acceptable for severe than mild behavior problems</td>
</tr>
<tr>
<td>Mittenberger, Parrish, Rickert, &amp; Kohr (in press)</td>
<td>Parents and grandparents receiving outpatient child behavior management services/100</td>
<td>DRO, RC, TO, spanking, medication</td>
<td>Cases described were 4-, 5-, or 6-year-old girls or boys with either tantrum, noncompliant, aggressive, or hyperactive behaviors</td>
<td>Treatment Evaluation Questionnaire (TEQ); a 12-item scale adapted from the TEI</td>
<td>(a) No differences between DRO, TO, and RC in acceptability, (b) spanking was least acceptable, (c) findings above were consistent across types of behavior problems, (d) medication for hyperactivity was less acceptable than behavioral procedures</td>
</tr>
</tbody>
</table>
The Treatment Evaluation Inventory includes 15 statements for evaluating a treatment's acceptability, each of which is rated on a 7-point Likert-type scale. The Semantic Differential contains a list of 20 statements about behavioral procedures that are rated on a 6-point Likert-type scale.

Collectively, the investigators examined several contextual variables for their impact on judgments of social acceptability, including (a) severity of the maladaptive behavior, (b) type of handicap of the individual, (c) reported effectiveness of the behavioral procedure, (d) adverse side effects, (e) rater characteristics (e.g., amount of teaching experience), (f) time required to implement the procedure, (g) risks of the procedure to others (e.g., classmates), (h) written versus videotaped case descriptions, and (i) effects of presenting educational information to the rater in the use of the procedures (see Table 3). No single study included investigation of all of these contextual variables.

The major findings pertinent to this review are presented in Table 3. A common finding was that behavioral procedures with positive reinforcement components (e.g., differential reinforcement of incompatible behavior) were judged more socially acceptable than others (e.g., time-out, overcorrection). These results were similar to those from previously reviewed studies (Augustine & Cipani, 1982; Irvin & Lundervold, 1988; Morgan & Striefel, 1987–1988). A relatively common finding was that more severe maladaptive behavior was associated with increased ratings of social acceptability of the procedure (Elliott et al., 1984; Frenz & Kelley, 1986; Martens et al., 1985; Miltenberger et al., 1989; Witt & Martens, 1983), although results from some studies were unsupportive (Kazdin, 1980b; Miltenberger et al., in press; Witt, Elliott, & Martens, 1984). Also, procedures with adverse side effects (Kazdin, 1981; Pickering & Morgan, 1985) and those that were time-consuming to implement (Elliott et al., 1984; Witt, Elliott, & Martens, 1984; Witt, Martens, & Elliott, 1984) were associated with low ratings of social acceptability.

Research on social acceptability has provided important information on judgments of behavioral procedures to decrease behavior. The findings indicate that judgments of social acceptability were often related to contextual variables. Unfortunately, comprehensive analysis of several relevant contextual variables has not been undertaken in a single study. Also, no investigators have examined the potential effects of other variables, such as expertise of the individual administering a procedure, length of time that an individual had experienced the maladaptive behavior, or duration or frequency of a procedure's use. Judgments of a procedure's frequency of use will be examined in the next section.

Some investigators have questioned the ecological validity of analogue research (Reimers et al., 1987; Witt & Robbins, 1985). That is, they have asked whether judgments in analogue situations are similar to or different from those in actual treatment intervention. Witt and Robbins (1985) called for examination of social acceptability by treatment consumers following actual implementation of behavioral procedures. Reimers et al. (1987) recommended additional research with treatment consumers to validate the findings of analogue studies.

A validity question related to the issue of analogue research methodology pertains to the respondents: who should judge the acceptability of procedures? Wolf (1978) argued that consumers who disliked treatment procedures would avoid them. Elliott (1986) stated that treatments judged unacceptable by consumers were, in a practical sense, no treatments at all. However, Kazdin (1977) cautioned that opinions of nonprofessionals, including treatment consumers, should not be viewed uncritically as prescriptive guidelines. He noted that consumers may not be familiar with empirical research on the use of behavioral procedures. Although in many studies on social acceptability teachers were used as raters, other professionals such as psychologists, program administrators, attorneys, or legislators, were not represented. Although not generally consumers of treatment, these and other professional groups would provide valuable data on the acceptability of behavioral procedures.

Given that behavioral procedures should be judged by treatment consumers after their actual implementation, another problem pertains to the number of procedures that a consumer can judge. Treatment consumers usually observe the implementation and effects of only one procedure. Unless more than one procedure is implemented, they cannot comparatively rate procedures' social acceptability. Consumer judgments of the social acceptability of multiple procedures have been examined in a few studies (Kirigin, Braukman, Atwater, & Wolf, 1982; Walle, Hobbs, & Caldwell, 1984), but some comparisons were of baseline-
treatment or treatment–no treatment conditions. Thus, more research is needed on judgments of social acceptability following actual implementation of multiple procedures.

Judgments of Frequency of Use of Procedure

Four studies were reviewed in which the estimated frequency of use of behavioral procedures was reported (Martens, Peterson, Witt, & Girone, 1976; SaUnd, Esquivel, & Baron-Pine, 1984; Wallace, Burger, Neal, van Brero, & Davis, 1976; Wood & Hill, 1983). Unlike most studies on judgments of social acceptability, survey respondents were professionals (e.g., regular and special education teachers) who were familiar with behavioral procedures. These respondents were asked to rate how often they used various procedures in their treatment setting. In some studies, respondents also rated aversiveness (i.e., unpleasantness: Wood & Hill), decision processes for determining usage (Wallace et al.), or judgments of effectiveness and ease of using the procedures (Martens et al.). Wallace et al. (1976) defined aversive procedures as the application of noxious stimuli that included electric shock, physical punishment, and chemical or auditory irritants. Although 80% of the respondents who were psychologists reported that aversive procedures should only be used as a last resort, these procedures were reportedly permissible in 45% of the institutions. Eleven percent of the respondents reported that aversive procedures had been used within the last 5 years but were discontinued. The procedures were used more often in institutions whose residents had severe or profound mental retardation. The latter finding is similar to more recent results from Guess et al. (1987), who reported that aversive procedures such as the presentation of punishing consequences (e.g., electric shock), negative reinforcement (i.e., avoidance of punishment), and overcorrection continue to be used in some institutions serving persons with severe or profound mental retardation. Interestingly, respondents from two institutions in the Wallace et al. (1976) study reported that legal action had been threatened by parents of residents if aversive procedures were discontinued. Both facilities agreed to continue using aversive procedures as the parents had insisted.

Wood and Hill (1983) examined the rated aversiveness and estimated frequency of use of 30 procedures. Aversiveness was defined as unpleasantness. Respondents were mostly regular and special education teachers. Results indicated that procedures rated as highly aversive (e.g., physical punishment) were infrequently used and that procedures not rated as aversive (e.g., verbal prompts, positive reinforcement delivered to other students) were more frequently used in classrooms. Similar results were obtained in large surveys of regular and special education teacher respondents (Martens et al., 1986; SaUnd et al., 1984). These results differed from those reported by Wallace et al. (1976) on the use of aversive procedures in institutional settings.

Summary and Conclusions

Collectively, findings from the studies reviewed revealed that: (a) procedures were ranked from least to most restrictive with general agreement; (b) the majority of respondents in the research agreed that procedures judged more restrictive should be used as a last resort; (c) more restrictive procedures were not frequently used in practice, except in some institutions for individuals with mental retardation; (d) respondents generally agreed that less restrictive procedures were more socially acceptable, and vice versa; and (e) the ratings of a procedure's social acceptability sometimes changed as a function of contextual variables such as the severity of the maladaptive behavior, adverse side effects, and time required for implementation.

Methodological issues were raised that may have limited the validity and generalizability of the studies' results. For example, some investigators did not define the dimension (e.g., restrictiveness) on which the behavioral procedures were judged. Others did not examine contextual variables (e.g., circumstances within which the behavior occurred, severity of the behavior to be decreased) that must always be considered in practice. None of the investigators considered contextual variables such as (a) the length of time that an individual had exhibited the maladaptive behavior, (b) the duration of use of a behavior procedure and the recipient's responsiveness, (c) the degree of expertise of the individual administering the procedure, or (d) variables that control a behavior's occurrence, such as an individual's attempts to communicate needs (Durand & Carr, 1985). Further, investigators did not consider the possible habituating effects on behavior when less to more restrictive procedures were implemented in sequence (Azrin, Holz, & Hake, 1963). Budd and Baer (1976) asserted that taking considerable
time to implement a series of less to more restrictive procedures may, in itself, represent a restrictive situation. Barton, Brulle, and Repp (1983) asked: "Are mildly aversive treatments applied over a long period of time more or less restrictive than highly aversive treatments that work quickly?" (p. 6).

Other methodological questions have been raised regarding the ecological validity of analogue research. Reimers et al. (1987) noted that analogue studies essentially involved judgments of acceptability prior to actual implementation of treatment, that is, respondents had no opportunity to observe the treatment, its effects on the behavior to be decreased, nor the outcome. Social acceptability judgments by consumers following the actual implementation of multiple procedures might validate the findings of these studies. Alternatively, differences between social acceptability ratings before and after actual implementation could be compared. Also, in addition to ratings by consumers of procedures' social acceptability, individuals with expertise in behavioral research should rate a procedure's acceptability.

Selecting procedures to decrease behavior should be based on empirical data regarding dimensions of effectiveness, restrictiveness, and social acceptability. Indeed, service providers need empirically based data to assist them in selecting procedures that are judged to be effective, acceptable, and least restrictive. However, considerable research must be conducted to provide sufficient data leading to the development of decision-making strategies that integrate these dimensions in a way that ensures valid selection and ethical treatment.

The controversy on procedures to decrease behavior has involved whether or not to prohibit use of aversive treatments. Contributing to this controversy has been the lack of a clear-cut, consensual definition of what constitutes aversive procedures (Matson & Taras, 1989; Snell, 1987). As often happens in public controversies, complex issues are reduced to two polarized positions, in this case, aversives versus nonaversives. The issues described in this review suggest that a for-or-against dichotomy is grossly oversimplified. The more critical and challenging issue seems to be that of identifying behavioral procedures that are most effective, most acceptable, and least restrictive in producing desired outcomes while preserving the individual's right to remain free from harm.

References


validation of decelerative (punishment) procedures by special educators of severely handicapped students. Research in Developmental Disabilities, 9, 331–350.


---

**Toward Independent Living & Learning (TILL), INC.’s**

**BEHAVIORAL TRAINING PACKAGE**

* AN ENTERTAINING AND INFORMATIVE MULTI-MEDIA PRESENTATION *
  Video Tapes, Workbooks, Handbooks, Out of Session Assignments, Pre and Post Tests, Staff Observation and Feedback Sheets

* COMPREHENSIVE *

Unit I: BEHAVIORAL PRINCIPLES—Cost $175
  Behavioral approach to treatment / Reinforcement / Teaching / Decreasing behaviors

Unit II: PROGRAM WRITING—Cost $125
  Behavioral assessment / Writing programs / Evaluating & revising

* EASY TO IMPLEMENT *

  Any supervisor can lead trainings using detailed Trainers’ Manual
  Can be done Individually, in small groups, or in a large group

* GUARANTEED TO PROMOTE LEARNING *

  Designed with errorless learning principles to ensure results
  Tested and proven effective with residential and day program staff

**BUY BOTH UNITS FOR ONLY $250**

TO ORDER, CONTACT: TILL, Inc. 133 East St., Dedham, MA 02026 (617) 329-6150