

# A 20 Year Review of Punishment and Alternative Methods to Treat Problem Behaviors in Developmentally Delayed Persons

Johnny L. Matson and Marie E. Taras

Louisiana State University

*Relevant journals were reviewed (n=23) for a 20 year period (1967 to 1987) to assess the status of treatments for severe behavior problems of developmentally delayed persons. A hand search of journals was made; 382 studies were identified. Procedures were analyzed by problem behaviors treated, side effects reported, whether the procedure involved painful stimuli, nonpainful stimuli, food satiation, positive procedures, extinction or combinations of methods. The number of studies reported yearly was also plotted. The implication of these data for federal and state policy makers and for treatment programs dealing with difficult to treat clients is discussed.*

Recently, the use of many effective behaviorally based treatments for developmentally disabled persons has come under close scrutiny. Some professionals and parent groups have argued that intrusive procedures should not be employed with handicapped persons. And, various groups have taken position statements opposing any aversive behavioral method. Among these groups are the Association for Persons with Severe Handicaps (TASH) and the Association for Retarded Citizens (ARC). Through the efforts of some who advocate the position of all positive procedures, some states (e.g., Massachusetts) have introduced bills to ban many behavioral procedures. More recently, positions supporting aversives along with reinforcement have been proposed by the Association for Behavior Analysis (ABA) and Mental

Retardation division of the American Psychological Association. Battle lines appear to have been drawn.

The debate has become perhaps the single most frequently discussed issue in the field of developmental disabilities and one of the most important topics in behavior modification. Also, the implications of this debate are likely to have a much broader impact in years to come in the treatment of schizophrenics, phobic and anxiety disorders, depression, weight control, and other behavior problems that are likely to employ intrusive procedures. Articles have recently appeared in the *New York Times*, *Boston Globe*, and *People Magazine*. The debate has been covered by "Niteline" and "20/20" a national ABC Program, and federal and state laws have been introduced. B. F. Skinner also has recently made a position statement supporting aversives since a misrepresentation of his views has been presented by some opposed to aversives (Landers, 1988).

The first issue to address is the definition of aversive or intrusive. It would appear based on the literature of those who favor a ban on aversive that the definition applies to the advocate perception versus the client. Thus, banning certain procedures is not based on whether they do or do not decelerate aberrant behavior, the traditional operant definition which Guess, Helmstetter, Turnbull, and Knowlton (1987) describe as a means-end approach. Rather, it appears that this position is based on what some see as inhumane procedures (Laski, 1987), although some disagreement exists among advocates as to what is and is not humane. Guess et al. (1987) equate lemon juice therapy (Fox, 1977; Reid, Tombaugh, & Heuvel, 1981; Sajuaj, Libet, & Agras, 1974) physical restraint and forced exercise, and other behavioral procedures with methods used to torture political prisoners (p. 24). The paper cited is by Amnesty International and entitled "Torture in the Eighties" (1984). Similarly, Turnbull (1986) in his presidential address states that aversive interventions such as punishments, negative reinforcement, and overcorrection are unwarranted. For the present study punishment will be defined as any stimulus that decreases the rate of behavior, while negative reinforcement is defined as the escape or avoidance of an unpleasant event resulting in the increase of a behavior since it results in the increase of appropriate behavior (Matson & McCartney, 1981). It is our view based on Turnbull's (1986) definition that some who oppose aversives are not familiar with standard definitions of some concepts (e.g., negative reinforcement).

It has been argued that aversives should not be employed for reasons other than being inhumane. It has been argued that: (a) there have been a small number of studies published over the last two decades; (b) many of the studies are inadequately designed; (c) effects occur only in the short term; (d) generalization is absent; (e) many negative side effects are evident; (f) the rate of research on methods that use an aversive component compared to

positive are not as striking as results with positives alone; (g) aversives result in increased in appropriate behavior and; (h) that the effects of the aversive treatments are not particularly strong (see Laski, 1987).

These are serious shortcomings which to date have not been supported by empirical data. The present study was designed to evaluate these issues. The present study was also designed to provide a more comprehensive review than previously published papers of this nature in scope and content (Gorman-Smith & Matson, 1985; Guess et al, 1987; Matson & Gorman-Smith, 1986). More journals over more years with more specific information were presented.

### *Evaluation Methods*

There were 23 journals searched including the *American Association for the Education of the Severely/Profoundly Handicapped Review*, *American Journal of Mental Deficiency*, *Behavior Modification*, *Behavior Therapy*, *Behaviour Research and Therapy*, *Child and Family Behavior Therapy*, *Clinical Psychology Review*, *Education and Training of the Mentally Retarded*, *Education and Treatment of Children*, *Journal of Abnormal Child Psychology*, *Journal of Applied Behavior Analysis*, *Journal of the Association for Persons with Severe Handicaps*, *Journal of Autism and Developmental Disorders*, *Journal of Behavior Therapy and Experimental Psychiatry*, *Journal of Child Psychology and Psychiatry*, *Journal of Clinical Child Psychology*, *Journal of Consulting and Clinical Psychology*, *Journal of Experimental Child Psychology*, *Journal of Mental Deficiency Research*, *Journal of Experimental Child Psychology*, *Journal of Special Education and Research in Developmental Disabilities*. A topic search using Psychological Abstracts, talking with experts in the field, and crosschecking references of published studies was the method employed to establish the list of studies. The goal was to produce the most comprehensive number of studies possible.

The text of each study was reviewed for the type of mentally retarded persons studied, their ages and intellectual level, various problems treated and procedures used. It was necessary to collapse across variables as a means of categorizing aspects of various papers. This method was employed to keep the number of variables at a manageable level. A brief review of some of the major decisions made follow.

The authors recognized that these decisions are to some degree arbitrary, but were our best guess based on the number and variations of the papers reviewed; (a) If several experiments were described in a paper, each experiment was tabulated separately; (b) Overcorrection studies were defined as those using positive practice, restitution, a combination of both methods or

forced arm exercises that were in the same topography as the problem behavior; (c) Time-out studies included those studies using removal from the environment, nonexclusionary time-out and withdrawal of reinforcement; (d) Studies containing persons from more than one age category were grouped in the mixed category; (e) Age categorization was based on age groups that occurred most frequently in the various studies; (f) Almost all the studies included reinforcement, therefore we categorized studies as punishment/aversive if both reinforcement and a punishment/aversive method were used, or a punishment/aversive method was used, or a punishment/aversive method was used alone; (g) When the effectiveness of two or more treatments were compared, the study was listed based on the most effective treatment; (h) For studies dealing with multiple target behaviors, it was possible to represent the study more than once (e.g., head banging would be listed under self-injurious behavior and body rocking would be listed as a stereotyped behavior); (i) Using the principle described in the previous point, if different treatments were used for different problem behaviors the study would be listed more than once; (j) Some studies used interesting and important methods, such as changes in bedtime, the use of corrective lenses, introduction of toys, variations of population density and changes in the room, and therefore were referred to as environmental/medical changes; (k) Since contingent and noncontingent exercise studies were few, they were grouped together; (l) A variety of behaviors where only one, two or three papers were found had been combined including tantrums, stealing, spitting, hyperventilation, food spillage, food refusal, out-of-seat, spoon banging, destroying property, noncompliance, and faking seizures; (m) Multiple behaviors were the combination of two or more problem responses, such as aggression and destroy; (n) Studies that used restraint as a reinforcer were listed as reinforcement; (o) Visual screening studies included both visual and facial screening, since there were not enough of each type of method to differentiate them; and (p) Relaxation training and response interruption were categorized as restraint.

### *Raters and Reliability*

Studies were reviewed by the second author, a Ph.D. student in clinical psychology with a master degree and two years of clinical experience. A second reviewer, an undergraduate psychology student, independently reviewed 11% of the same studies as a reliability check. The agreement on the categories listed above by the authors averaged 90% using the standard percent agreement formula. Reliabilities ranged from 83% (side effects) to 100% (age category).

### Findings

A variety of factors were evaluated. One factor that deserves consideration is the number of studies being conducted. Advocates have claimed that a decrease in aversive studies has occurred because such methods are not particularly effective. The rate of all studies for the last 7 years which deal primarily with problem behaviors however have been stable, with most articles employing an aversive component (see Figure 1). These findings are in contrast to previous claims. Given the recent increase in public attention on the topic via newspaper, television, and professional reports it is hoped that more treatment studies will be forthcoming, particularly those of a positive nature. However, it should be cautioned that an increase or decrease in positive versus aversive treatment procedures in studies does not appear to indicate overall effectiveness of procedures. This hypothesis is a directly testable one and no assumptions should be made until these direct comparison studies have been published.

The number of studies on this topic was greater than with any other target behavior in the developmental disabilities area. This finding also is in

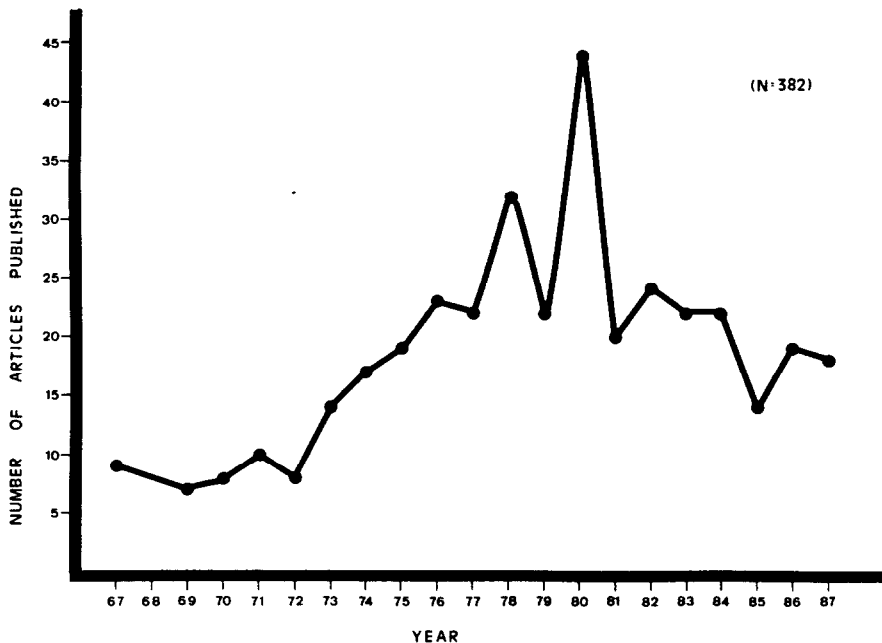


FIGURE 1. Number of publications utilizing decelerative procedures on aberrant behavior by year published.

contrast to criticism made regarding aversives as was our finding regarding methodological design. Using criteria on single case designs such as baseline data, reliability, and other basic methodological controls accepted as standard for single case research (Barlow & Hersen, 1984) these studies were as methodologically sound as any other area of developmental disabilities. To hold these studies to a higher methodological standard would be out of line with existing research methodology norms and would in essence invalidate studies of equal or greater proportion on positive procedures to treat not only the most severe problem behaviors but all problem behaviors of the developmentally disabled.

The short term nature of effects has also been cited as a reason to employ reinforcement versus punishment plus reinforcement or punishment alone. We define follow-up as occurring at least one month after treatment was terminated so that a relatively stringent criteria with respect to treatment research in general could be established. We then calculated the percentage of studies by treatment category that had follow-up data of this type. Percentages were: combined aversives (52), nonpainful stimuli (46), painful stimuli (43), extinction (36), positive procedures (26). All studies reporting follow-up showed significant maintenance of treatment gains. Maintenance of effects did not significantly differ although the least amount of follow-up was available for the positive procedures. Based on these data, the conclusion that positive methods result in greater long term effects than aversives and that short term effects only occur with aversives does not seem to be supported by the data.

It has been argued that the effects obtained with aversive procedures are not substantial. Once again we had no information which would support such a contention. Effects were typically quite rapid. It should also be pointed out that these studies would most likely have been unacceptable to journal reviewers and would not have been published had the treatment effects not been pronounced.

Generalization has also been questioned with aversives. Our review suggests a clear need for more and better generalization data with behavior problems on developmentally disabled persons. This need extends however beyond any one technique. Our findings suggest a need for considerably more information or generalization effects with all the procedures studied to date. However, it should also be pointed out that the demonstration that such techniques work is a major and only recent development. Generalization research is needed but would be a logical next step on research with existing technologies. This research along with direct comparison studies of various treatment procedures is now needed.

Another issue of concern, represented by Table 1, was the breakdown of treatment procedures by age for persons studied. A considerable amount of information is available in the table. However, some general findings may be

**TABLE 1.**  
**Breakdown of Studies by Treatment Method and Age**

TRT Method	0-10 yrs.	11-15 yrs.	16+ yrs.	Mixed Ages	Missing
<i>Aversives:</i>					
<i>Painful stimuli:</i>					
electric shock	10	6	9	2	
ammonia	10	1	2		
taste	7	3	1		1
slap	4		1	1	
icing	1				
cold bath	1				
<i>Nonpainful stimuli:</i>					
overcorrection	40	10	41	5	1
timeout (TO)	27	11	14	8	1
restraint	14	4	12	7	
visual screening	9	5	4		1
exercise	7		3	5	
reprimand	9		2	2	
water mist	3	2	2	1	
response cost	3	2	3	1	
tickling		2			
white noise	1				
<i>Food satiation:</i>					
			3	1	
<i>Extinction:</i>					
Extinction (ignore)		3	1	2	1
Sensory extinction	3	1		3	
<i>Positive procedures:</i>					
reinforcement	10	2	11	7	
differential reinforcement	9	10	10	3	1
increased activities	3	1	3	5	
decreased demands	5	1		3	
environ./med. changes	7	1	2	4	
self-monitoring		1	3		
functional communication	1			2	
physical therapy	1				
task variation			1	1	
instr., feedbk., modeling			2		
<i>Combined procedures:</i>					
TO + restraint	7	5	5		
TO + slap	1				
TO + spank		1			
TO + overcorrection			2	1	
taste + TO	1		1		
taste + overcorrection	1				
taste + TO + restraint			1		
shock + TO			2		
food satiation + TO			1		
food satiation + overcorr.			1		
satiation + overcorrection			1		
water mist + screening			2		
water mist + overcorrection			1		
slap + restraint		1			
increased activities + TO			1		
response cost + func. comm.			1		
response cost + restraint	1				
response cost + extinction		1			
hair-tug, shock, + restraints					1

Note: trt = treatment; yrs. = years; environ. = environmental; med. = medical; instr. = instructions; feedbk. = feedback; func. = functional; comm = communication.

of particular note. Studies that employ painful stimuli have largely been used with children under 10 years of age. Advocates opposed to such procedures are likely to be particularly troubled by this information. Why does this pattern emerge? It is likely that some of these methods are less practical with older clients (e.g., electric shock equipment would more likely be torn off by older clients), and if the treatment was effective there would be less need for it with older clients. Secondly, many problems reviewed are very destructive and potentially harmful to self and others. Therefore, they must be treated as soon as they appear to head off injury. Furthermore, traditionally there has been a focus on prevention of severe behavior problems with young persons among behavior modifiers, other clinicians and educators. The bulk of all the treatment research to date is on younger children. Thus, these data are merely a reflection of research on treatment with the developmentally disabled in general. Finally, for genetic and environmental reasons, shorter lives have in the past been common place for handicapped persons relative to the general population. Thus, few studies of any type with aged developmentally disabled persons may be reflected to some degree in the disproportionate number of young persons with handicaps who have been studied.

A review of how frequently various treatment procedures have been studied is also an issue of some importance. It has and should be argued that procedures with the best data base should be a primary consideration in determining treatment of choice. Based on this factor the most thoroughly researched painful stimulus is contingent electric shock. Advocates have recently argued that this method should never be used and that it is in disfavor because few studies have appeared using this method. Also, it has been stated that now there exists more effective alternatives. A more plausible explanation is that many organizations have become concerned about using these methods because of the advocate push to make them controversial. Also, the fact that such methods have proven so effective may suggest the need for less research on this procedure, relative to other behavioral treatments. Finally, the misperception should not be left that research is no longer being done on this technique. Foxx has a four follow-up study in press with the *American Journal of Mental Retardation* and Lincheid is in the process of submitting research on a new shock device called SIBIS. Data by the latter group have already appeared at professional meetings.

The most frequently employed nonpainful stimuli is overcorrection. It is noteworthy that contingent electric shock, and overcorrection are among the methods most frequently criticized by advocates (Turnbull, 1986). There has been a drop off on research using this technique. It is unclear whether this is due to the factors noted above and/or for other reasons.

Differential Reinforcement of Other Behavior (DRO) is the most frequently employed of the positive procedures. Insufficient research has been done on other positive treatments such as functional analysis of behavior to



draw firm conclusions. It is likely given current trends that within five years scientists and clinicians will have a firmer understanding of how effective these positive methods will be with extreme problem behaviors. We are of the opinion that these methods will prove effective with some behaviors currently treated with aversives, but not all.

A final breakdown of studies was done by target behavior. Three categories were selected, including self-injury, stereotypies, and disruptive and aggressive behavior. Self-injury problems that appeared most frequently were multiple self-injurious behaviors, biting, and head hitting. And, while a range of procedures were employed, contingent electric shock and overcorrection were the most commonly used methods. Differential reinforcement proved to be the most common of the positive methods with 12 studies compared to 34 with overcorrection. There were a total of two differential reinforcer studies for multiple self-injury, biting, and head hitting, and 14 using overcorrection. It is not suggested however that overcorrection should be routinely selected over differential reinforcement (Matson & DiLorenzo, 1984). Nevertheless, some have strongly advocated for using reinforcement only for such problems (LaVigna & Donnellan, 1986). The available data does not, at this point, seem to support such a contention (Bailey, 1987). Hopefully, research in the next few years will support a closer approximation of this hypothesis of positive procedures only. Much more research is needed before such conclusions could be made, however.

Table 2 presents information on treatments for self-injurious behaviors. Overcorrection followed by contingent electric shock has the best data base, with overcorrection having been used over a broader range of target behaviors. Several other common procedures for such problems include time-out, restraint, visual screening, differential reinforcement, and increased activities.

Table 3 shows that for disruptive/aggressive behavior, electric shock was again the most commonly employed painful stimuli. However, the most common nonpainful treatments were overcorrection and time-out. The most widely used positive procedures were reinforcement and differential reinforcement. Yet, overcorrection and time-out were still the predominantly implemented treatments for disruptive and aggressive behavior.

Research with stereotypies is reflected in Table 4. As would be expected, painful stimuli were rarely reported. There was one contingent electric shock study. This data suggests that researchers and clinicians have for many years been reluctant to use highly intrusive treatments with less severe problem behaviors. These data then seem to confirm the use of painful stimuli for the more severe behavior problems and is in line with what most state regulatory bodies and practitioners recommend. Similarly, there was a much lengthier list of positive procedures for stereotypies as compared to self-injury.

Several issues in addition to the specific information presented are also

**TABLE 2.**  
**Breakdown of Studies by Target Behavior and Treatment Method for Self-Injurious Behavior (SIB)**

TRT Method:	Head-banging	Ruminat./vomit.	Hair-pulling	Biting	Head-hitting	Pica	Picking	Body-slapping	Eye-poking	Multiple SIB
<i>Combined procedures:</i>										
TO + restraint						1				6
TO + spank										1
TO + overcorrection										1
taste + TO		1								
taste + overcorrection		1								
taste + TO + restraint										
shock + TO										1
food satiation + TO		1								
food satiation + overcorrect.		1								
water mist + screening					3					
water mist + overcorrection					3					
restraint + extinction				1						
hair-tug, shock, + restraints										1
<i>Aversives:</i>										
<i>Painful stimuli:</i>										
electric shock	3	8		1	3					4
ammonia			3	1	2	1			1	1
taste		5		2						1
icing						1				

<b>Nonpainful stimuli:</b>										
overcorrection	4	3	2	4	7	6	1	1	3	3
timeout (TO)	1	2		2	4		1	1	1	8
restraint	1			2	4	4			1	7
visual screening			1	2	2	1				1
exercise		1		1			1			3
reprimand	1								1	
water mist				2		1				1
response cost				1						1
tickling	1									
<b>Food satiation:</b>										
Extinction		4								
<b>Extinction (ignore)</b>										
Sensory extinction					2					1
<b>Positive procedures:</b>										
reinforcement		1		1	1					4
differential reinforce.	4	4		1		1	1			1
increased activities						1			1	3
decreased demands										1
environ./med. changes		2		1	1					1
self-monitoring							1			1
functional communication										1

*Note.* trt = treatment; ruminat. = rumination; vomit. = vomiting; reinforce. = reinforcement; environ. = environmental; med. = medical

**TABLE 3.**  
**Breakdown of Studies by Target Behavior and Treatment Method for Disruptive and Aggressive Behavior**

TRT Method:	Rocking	Mouthing	Toe-walking	Hand play	Tongue protrusion	Head movts.	Hands-to-head	Rituals	Multiple stereoty.
<i>Aversives:</i>									
Painful stimuli:									
electric shock	1								
ammonia				1					
slap	1			1					3
Nonpainful stimuli:									
overcorrection	1	7	1	8	1	3	1	1	18
timeout (TO)		1		5			2		3
restraint	1	3		2	1				6
visual screening		2						1	2
exercise		1							5
reprimand	1	1							2
water mist		1		1				1	
white noise									1
<i>Extinction:</i>									
sensory extinction				1					4
<i>Positive procedures:</i>									
reinforcement	1	1					1	1	4
differential reinforce.		1							4
increased activities		1		1					2
decreased demands		1							
environ./med. changes	2			1					7
self-monitoring					1	1			
functional communication									1
physical therapy									1
task variation				1					1
<i>Combined procedures:</i>									
TO + restraint		1		1					3
TO + slap									1
water mist + screening		1							
slap + restraint									1

*Note:* trt = treatment; stereoty. = stereotypies; environ. = environmental; med. = medical

TABLE 4.  
Breakdown of Studies by Target Behavior and Treatment Method for Stereotypes

TRT Method:	Aggression towards others	Inapprop. vocalizat.	Public disrobing	Enuresis/ encopresis	Public masturbation	Disruptive behavior	Multiple behaviors
<i>Aversives:</i>							
Painful stimuli:							
electric shock	2	2				2	1
ammonia	1						
taste	2				1	1	
slap						1	
cold bath				1			
Nonpainful stimuli:							
overcorrection	4	3	3	3	1	13	4
timeout (TO)	4	11	1			14	8
restraint	1				1	2	2
visual screening		3			1	2	
exercise	2	1				1	2
reprimand	1	3				2	
water mist	1						
response cost	1	2				2	2
tickling							1
<i>Extinction:</i>							
Extinction (ignore)	1	1				1	
Sensory extinction		1				1	
<i>Positive procedures:</i>							
reinforcement	4	3		1		8	3
differential reinforcement	2	5	1			7	5
increased activities						3	
decreased demands	2	1				2	1
environ./med. changes						2	1
self-monitoring						1	
functional communication		1					1
task variation							2
instr., feedbk, modeling		2					

(continued)

TABLE 4.  
Continued

TRT Method:	Aggression towards others	Inappropri- vocalizat.	Public disrobing	Enuresis/ encopresis	Public masturbation	Disruptive behavior	Multiple behaviors
<i>Combined procedures:</i>							
TO + restraint	3					4	2
TO + slap	1						
TO + spank	1						
TO + overcorrection				2			
taste, TO + restraint	1						
shock + TO							1
satiation + overcorrection						1	
increased activities + TO						1	
response cost + func. comm.						1	
response cost + restraint							1

Note. trt = treatment; inapprop. = inappropriate; vocalizat. = vocalization; instr. = instructions; feedback = feedback; func. = functional; comm. = communication

worth noting in this national debate, related to these data. A perplexing point is that many advocates and national groups, such as TASH have taken the position that punishment procedures are not acceptable to the general population. This question is however a very complex one since the clients, their age, the severity of their behavior problems, and parental beliefs, may effect such decisions in a very idiosyncratic way. Our position is that such issues, given the potential magnitude of change in availability of treatments being suggested, warrant and deserve empirical investigation. Frentz and Kelley (1986), for example, found in a direct test of treatment acceptability that parents found response cost, a punishment procedure, to be far more desirable as a treatment for their children's noncompliant behavior than differential reinforcement. Several reasons, such as the effectiveness of the methods and the lack of difficulty in implementing them, led to these preferences. Also, the nature of the problem is at issue. Parents whose child had severe life threatening self-injurious behavior which had been ineffectively treated with a number of procedures would be likely to accept more intrusive methods than a parent of a young handicapped child with little or no history of maladaptive behavior. Various parameters are likely to result in various choices if past research on this topic is a predictor. It would seem from these and related findings that much needs to be done to differentiate fact from folklore. Our view is that case by case reviews rather than all-or-none national policies will likely result in a more flexible and acceptable model for treatment choice and implementation.

Guess et al. (1987) argue for better regulation of punishment procedures. We strongly agree and have previously laid out guidelines on how this might be done (Matson & Kazdin, 1981). It is our contention that such regulation has been effective with other disciplines. We would not wish to have a nurse practitioner perform open heart surgery. Nor would we wish to see surgery banned. The issue then should be proper regulation and peer review to insure that where very severe problems exist (e.g., biting off fingers, blinding oneself) that treatments to curb such problems including punishment methods be available and implemented by trained professionals. This, in our view, is a moral issue. Parent or patient consent, properly qualified professionals, and peer reviews should be included with aversive being a last resort, *but an available one*.

Misrepresentation or misunderstanding of procedures also seems to be evident in the literature. McGee and Hobbs (1987) have recently listed various procedures used to treat developmentally disabled clients. Examples listed include squirting ammonia in a person's face, beatings, isolation, and tying persons like animals to metal frame beds. There were no published studies using these methods. It would seem then that all could agree that these procedures are not accepted practice. Also, it should be clearly noted that these are not behavior modification techniques. Other procedures em-

ployed in scientific studies have also been described by McGee and Hobbs (1987) as torturous and repugnant such as tickling, contingent electric shock, and squirting a fine water mist into a client's face for aggression. As noted previously, it is our view that a more systematic procedure might be possible to draw conclusions regarding the treatment acceptability of procedures (Kazdin, 1980; Parsons, Schepis, Reid, McCarn, & Green 1987). Using a well developed methodology of this sort should help clinical psychologists and other health professionals understand what treatments are accepted by whom and under what conditions.

A frequent criticism of punishment/aversive methods is that many negative side effects result. This position has been a frequent justification to discontinue these treatment methods. An assumption of this type may be based on the early animal literature (Arzin & Holtz, 1966). However, the data obtained in our review of 382 applied studies does not support this contention. Table 5 provides example behaviors which were antedotally reported in most studies where a side effect was noted. It could be contended that authors may have emphasized the more positive aspects of their studies. However, the studies reviewed were by and large evaluated prior to publication by independent peer reviewers and most of the studies were published before the current controversy over aversives developed. An interesting and striking development was the number of positive side effects ( $n=212$ ) to negative side effects ( $n=16$ ) reported regardless of treatment. This rate is 93% positive for all side effects reported. Also, there are rather small differences in side effects by treatment. Time-out and restraint ( $n=5$ ), DRO ( $n=12$ ) and increased activities ( $n=4$ ) resulted in 100% positive side effects. However, contingent electric shock ( $n=56$ ), physical restraint ( $n=24$ ), and combined punishment procedures such as time-out plus restraint, and time-out plus contingent slaps ( $n=25$ ) resulted in a 96% rate of positive side effects. Given that increased numbers of side effects are likely to result simply by the larger number of studies on aversive procedures. The differences in treatment with 100% positive side effects ( $n=21$ ) and with ( $n=104$ ) positive effects 96% may be largely due to chance. This point is further enhanced by our finding that time-out and restraint demonstrated positive effect rates equal to DRO. Also, the severity of side effects were relatively mild compared to the target behaviors treated in that they did not lead to injury of self or others, the typical target behaviors for which the interventions were used.

It has been argued that painful stimuli in particular and aversives in general are inhumane, are unethical and immoral and that the absence of negative side effects cannot be confidently asserted and certainly not assumed (Laski, 1987). These data clearly show that more intrusive procedures result in roughly equivalent rates of positive to negative side effects when compared to positive procedures. Data of this sort should be viewed cau-



**TABLE 5.**  
**Observed Side Effects of Decelerative Procedures**

Sample of Side Effects		
TRT & Behavior	Positive	Negative
<i>Aversives:</i>		
Painful stimuli:		
<i>electric shock:</i>		
– rumination	Weight gain (7) Increased social behavior (5) Decreased crying & tantrums Improve self-feeding Increased activity levels	
– head hitting	Decreased whining	Increased collateral behav.
– head banging	Decreased crying	
– multiple SIB	Removal of restraints	
– disrupt. behav.	Increased eye contact Improved interactions Greater attentiveness	
<i>taste</i>		
– biting	Decreased disruptions Wounds healed Removal of splints	
– multiple SIB	Increased social participation	Increased collateral behav.
– rumination	Weight gain (3) Increased social & comm. skills Increased social behavior (3) Decreased crying More responsive to environment	
<i>slap</i>		
– rocking	Decreased 2 collateral behaviors Increased academic skills	
– hand play	Increased academic skills	Decreased social behavior Increased aggression
– multiple stereoty.		
<i>Nonpainful stimuli:</i>		
<i>overcorrection:</i>		
– hands-to-head	Decreased 2 collateral behaviors	Increased stereotypies
– hand plays	Increased sociability (2)	
– multiple stereoty.	Decreased collateral behavior Wounds healed More alert & aware of environment Increased approp. toy play (3)	Increased stereotypies

(continued)

tiously. However, previous assumptions have been totally speculative. Until other data based information is available it would seem reasonable to accept the information presented here.

Critiques have also suggested that painful stimuli are inhumane, unethical, and immoral. One must determine on what grounds such assumptions lie. It has been demonstrated here that marked differences in side effects do

**TABLE 5.**  
**Continued**

Sample of Side Effects		
TRT & Behavior	Positive	Negative
— head hitting	Wounds healed Decreased 2 collateral behaviors Removal of restraints	Increased stereotypies
— head-banging	Increased social participation	Increased SIB
— pica	Increased sociability Increased sociability (3) Elimination of whipworm (2) Removal of restraints Improvements in teeth & gums Learned new behavior (2)	
— hands-to-head		Increased collateral behav.
— hand play	Decreased collateral behaviors (2)	
— aggression	Faded restraints Increased sociability	
— disrupt. behav.	Faded restraints Learned new behaviors Increased time spent in approp. behav.	
— inapprop. vocaliz.	Increased sociability	
— multiple disrupt.	Increased sociability	
— hand play & disrupt. behav.	Decreased collateral behavior	
<i>restraint</i>		
— head-hitting	Increased compliance	
— pica	Decreased collateral behavior	
— disrupt. behav.	Increased sociability Increased academic work	
— multiple disrupt.	Increased physical contact	
<i>visual screening</i>		
— pica	Increased spontaneous play	
— hair pulling	Wounds healed	
— multiple SIB	Decreased collateral behaviors	
— head-hitting	Wounds healed	
— disrupt. behav.	Increased sociability	
<i>exercise</i>		
— multipel SIB	Increased sociability (2) Decreased 2 collateral behaviors	
— biting	Increased sociability Decreased collateral behavior Increased academic progress	
— multiple stereoty.	Increased approp. play Increased academic responding	

not appear to be greatly different across procedures. Also, long term effects have been demonstrated in some studies such as the four year follow-up of a continient shock program recently reported by Foxx, and the approximately 40% rates of all punishment studies reviewed.

Another criticism is the statement that a small number of studies have been conducted, typically with poor methodological control (Laski, 1987). More studies have been published in this area than any other with developmental disabilities in the last 20 years. Methodological criteria were similar for positive and aversive techniques. On what criteria then are inhumane, immoral, and unethical judged? There are no factually based data on which to draw such a conclusion. Perhaps the data collected here will result in a more systematic means of establishing the value of various behavioral treatments. Additionally, systematic study of treatment acceptability may lead to a more accurate evaluation of these various treatments. It would appear that all professionals would agree that the most humane, effect treatment should always be used. We believe such an important decision should be factually based.

## REFERENCES

- Azrin, N. H., & Holtz, W. C. (1966). Punishment. In W. K. Honig (Ed.), *Operant behavior: Areas of research and application*. New York: Appleton-Century-Crofts.
- Bailey, J. S. (1987). Misguided alternatives. *Contemporary Psychology*, *32*, 571-572.
- Barlow, D. H., & Hersen, M. (1984). *Single case experimental designs: Strategies for studying behavior change*. New York: Pergamon Press.
- Chafee (1987). Medicaid Home and Community Quality Services Act of 1987. Amendment to Title 19 of Social Security Act.
- Foxx, R. M. (1977). Attention training: The use of overcorrection avoidance to increase the eye contact of autistic and retarded children. *Journal of Applied Behavior Analysis*, *10*, 489-499.
- Frentz, C., & Kelley, M. L. (1986). Parents acceptance of reductive treatment methods: The influence of problem severity and perception of child behavior. *Behavior Therapy*, *17*, 75-81.
- Gorman-Smith, D., & Matson, J. L. (1985). A review of treatment research for self-injurious and stereotyped responding. *Journal of Mental Deficiency Research*, *29*, 295-308.
- Guess, D., Helmstetter, E., Turnbull, H. R., & Knowlton, S. (1987). *Use of aversive procedures with persons who are disabled: An historical review and critical analysis*. Seattle, WA: The Association for Persons with Severe Handicaps.
- Kazdin, A. E. (1980). Acceptability of alternative treatments for deviant child behavior. *Journal of Applied Behavior Analysis*, *13*, 259-273.
- Landers, S. (1988). Skinner joins aversive debate. *American Psychological Monitor*, *19*, 22-23.
- Laski, F. (1987). Epilogue. In D. Guess, E. Helmstetter, H. R. Turnbull, & S. Knowlton (Eds.), *Use of aversive procedures with persons who are disabled: An historical review and critical analysis*. Seattle, WA: The Association for Persons with Severe Handicaps.
- LaVigna, G. W., & Donnellan, A. M. (1986). *Alternatives to punishment: Solving behavior problems with non-aversive strategies*. New York: Irvington Publishers, Inc.
- Matson, J. L., & DiLorenzo, T., (1984). *Punishment and its alternatives. A new perspective for behavior modification*. New York: Springer.
- Matson, J. L., Gorman-Smith, D. (1986). A review of treatment research for aggressive and disruptive behavior in the mentally retarded. *Applied Research in Mental Retardation*, *7*, 95-103.

- Matson, J. L., & Kazdin, A. E. (1981). Punishment in behavior modification: Pragmatic, ethical and legal issues. *Clinical Psychology Review*, *1*, 197-210.
- Matson, J. L., & McCartney, J. R. (1981). *Handbook of behavior modification with the mentally retarded*. New York: Plenum Press.
- McGee, J. J., & Hobbs, D. C. (1987). *Gentle teaching: A non-aversive approach to helping persons with mental retardation*. New York: Human Sciences Press.
- Parsons, M. B., Schepis, M. M., Reid, D. H., McCarn, J. E., & Green, C. W. (1987). Expanding the impact of behavioral staff management: A large-scale long-term application in schools serving severely handicapped students. *Journal of Applied Behavior Analysis*, *20*, 139-150.
- Reid, G., Tombaugh, T., & Hauvel, K. (1981). Application of contingent physical restraint to suppress stereotyped body rocking of profoundly mentally retarded persons. *American Journal of Mental Deficiency*, *86*, 78-85.
- Sajuaj, T., Libet, J., & Agras, S. (1974). Lemon juice therapy: The control of life-threatening rumination in a six-month old infant. *Journal of Applied Behavior Analysis*, *7*, 557-563.
- Turnbull, H. R. (1986). Presidential Address 1986: Public policy and professional behavior. *Mental Retardation*, *24*, 265-276.