Superimposition and Withdrawal of Tangible Consequences as a Treatment for Automatically Reinforced Problem Behaviors

Tina M. Sidener
Western Michigan University

Follow this and additional works at: https://scholarworks.wmich.edu/masters_theses

Part of the Applied Behavior Analysis Commons

Recommended Citation
https://scholarworks.wmich.edu/masters_theses/4609
SUPERIMPOSITION AND WITHDRAWAL OF TANGIBLE CONSEQUENCES AS A TREATMENT FOR AUTOMATICALLY REINFORCED PROBLEM BEHAVIOR

by

Tina M. Sidener

A Thesis
Submitted to the
Faculty of The Graduate College
in partial fulfillment of the
requirements for the
Degree of Master of Arts
Department of Psychology

Western Michigan University
Kalamazoo, Michigan
June 2004
ACKNOWLEDGMENTS

I would like to begin by thanking my advisor, Dr. James Carr, for his constant encouragement and support, and for taking the time to shape my behavior as a researcher. It has truly been an honor to study with him, and his enthusiasm for research will continue to inspire me throughout my career.

Secondly, I would like to thank my committee members and mentors, Dr. Linda LeBlanc and Dr. Jack Michael, for taking time to review my work, and for the excellent teaching and supervision I have received from them throughout my graduate career.

Lastly, I would like to thank my husband, Dave, for introducing me to behavior analysis and for his support and perseverance as we both pursued graduate degrees.

Tina M. Sidener
SUPERIMPOSITION AND WITHDRAWAL OF TANGIBLE CONSEQUENCES AS A TREATMENT FOR AUTOMATICALLY REINFORCED PROBLEM BEHAVIOR

Tina M. Sidener, M.A.
Western Michigan University, 2004

Tangible superimposition and withdrawal is a reductive procedure in which a new stimulus is delivered following behavior already maintained by a different controlling stimulus. The new stimulus is then removed in an attempt to reduce behavior. The current investigation sought to extend previous research on this procedure by evaluating its efficacy and durability as a treatment for stereotypy in three children diagnosed with autism. First, automatic reinforcement functions for stereotypic behaviors were identified via functional analyses. Next, for two participants, tangible items were delivered contingent upon stereotypy and then subsequently withdrawn. When the superimposition procedure proved ineffective, environmental enrichment was implemented and was found to be efficacious in reducing the stereotypy of both participants.
# TABLE OF CONTENTS

ACKNOWLEDGEMENTS .............................................................................. ii

LIST OF FIGURES ...................................................................................... v

INTRODUCTION ........................................................................................... 1
    Defining Features .................................................................................. 1
    Etiology ............................................................................................... 2
    Rationale for the Treatment of Stereotypy ............................................. 3
    Behavioral Interventions ..................................................................... 4

METHOD ....................................................................................................... 21
    Participants ......................................................................................... 21
    Dependent Variables and Data Collection .......................................... 22
    Informant and Direct-Observation Assessment .................................... 23
    Skills Assessment ............................................................................... 24
    Functional Analysis ............................................................................ 25
    Interobserver Agreement .................................................................... 28
    Treatment Evaluation .......................................................................... 28
    Environmental Enrichment ............................................................... 31
    Independent Variable Integrity ......................................................... 32

RESULTS ..................................................................................................... 34
    Functional Analysis ............................................................................ 34
    Treatment Evaluation ......................................................................... 37
Table of Contents—Continued

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>DISCUSSION</td>
<td>41</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>47</td>
</tr>
<tr>
<td>APPENDIX</td>
<td>60</td>
</tr>
<tr>
<td>Approval Letter From the Human Subjects</td>
<td>61</td>
</tr>
<tr>
<td>Institutional Review Board</td>
<td></td>
</tr>
</tbody>
</table>
LIST OF FIGURES

1. Functional Analysis Outcomes for Each Participant ........................................ 36
2. Treatment Evaluation Outcomes for Rose and Libby .................................... 38
3. Free-operant Preference Assessment Outcomes for Rose and Libby ................. 40
INTRODUCTION

For more than 50 years, researchers have sought to develop effective reductive techniques for the treatment of stereotypic behaviors (e.g., Berkson & Mason, 1964; Lourie, 1949). However, much controversy has surrounded the terminology, defining features, etiology, and treatment of these behaviors. For example, in different lines of research, stereotypy has also been termed self-stimulation, autistic mannerisms, disturbances in motility, bizarre gestures, and stereotyped behavior (e.g., Koegel & Koegel, 1989; Rojahn & Sisson, 1990; Ross, Yu, & Kropla, 1998; Weisberg, Passman, & Russell, 1973). In the current manuscript, the term stereotypy or stereotypic behavior will be used to avoid an inference of behavioral function.

Defining Features

Part of the difficulty in defining stereotypy is that it does not constitute a specific topography or response class, but rather a manner in which a variety of behaviors may be performed (King & Krishnamoorthy, 1998). In fact, in a review of 60 treatment studies, LaGrow and Repp (1984) identified 50 topographies of stereotypy, the most common being body rocking, mouthing, complex hand and finger movements, thumb/limb sucking, and limb/body posturing. Although integral to a diagnosis of the Pervasive Developmental Disorders (PDD), stereotypy may warrant a distinct diagnosis of Stereotypic Movement Disorder when PDD characteristics are not manifested. The Diagnostic and Statistical Manual of Mental Disorders (4th ed.; DSM-IV) defines these movements as “repetitive, seemingly driven, and nonfunctional motor behaviors” (American Psychiatric Association, 1994, p. 134). Other definitions have been suggested, most describing the repetitive, highly consistent nature of stereotypy, as well
as reflecting the authors' conceptual approach concerning the controlling stimuli and adaptive value of these behaviors. For example, while some researchers have defined stereotypic behaviors in terms of unknown controlling variables (e.g., Rollings, Baumeister, & Baumeister, 1977; Schroeder, 1970), others have classified them as behaviors controlled by sensory or automatic reinforcers (e.g., Lovaas, Newsom, & Hickman, 1987). Automatic reinforcement refers to the maintenance of behavior by consequences that are not socially mediated, but rather produced by the behavior itself (Vaughan & Michael, 1982). Before the development of functional analysis technology (Iwata, Dorsey, Slifer, Bauman, & Richman, 1994/1982), stereotypy was often assumed to be automatically reinforced (e.g., Lovaas et al., 1987). However, recent advances in functional assessment have made possible the identification of social functions of stereotypy as well (e.g., Goh et al., 1995; Kennedy, Meyer, Knowles, & Shukla, 2000; Mace & Belfiore, 1990). Because a central characteristic of stereotypy seems to be that a social function is not apparent (Baumeister & Forehand, 1973; LaGrow & Repp, 1984), an appropriate definition for stereotypy may be any topography of behavior that is performed in a highly consistent and repetitive or prolonged manner and for which a specific social function is not evident.

Etiology

Although typically developing children engage in behaviors that appear similar to stereotypy, it is not clear why individuals with some disorders continue to do so at such high rates (Berkson, 1996). Etiological theories of stereotypy have been proposed from psychodynamic, physiological, neurochemical, and behavioral perspectives. Psychodynamic approaches, which have attributed stereotypy to efforts at fulfilling needs
created by an impaired parent-child relationship, remain without scientific basis (O'Brien, 1981). One of the main physiological theories states that the performance of stereotypy allows for homeostatic balance in the central nervous system when environmental stimuli become over- or under-stimulating (Rojahn & Sisson, 1990). Although plausible, this theory is difficult to prove or disprove due to the tautological nature of its basic logic. Neurochemical theory posits that the function of neurotransmitters such as dopamine may be impaired in some way in individuals who exhibit high rates of stereotypy (Lewis, Baumeister, & Mailman, 1987). However, changes in several different neurotransmitters have been associated with stereotypy, and “from neurochemical, neuroanatomical, and neurofunctional vantage points, ... stereotypic behaviors are quite heterogeneous and rather poorly understood” (Grossman & Verobyev, 1998, p. 98). Finally, behavioral theory holds that stereotypy is maintained operantly or by schedule induction. Although there has been some evidence for an adjunctive etiology (e.g., Lerman, Iwata, Zarcone, & Ringdahl, 1994), most behavioral research has been based on the operant control of stereotypy by automatic or social reinforcers. Treatments based on behavioral theory are common and have been reported as relatively effective (Rojahn & Sisson, 1990).

Rationale for the Treatment of Stereotypy

Research has supported the importance of treating high-rate stereotypy. The presence of these behaviors alone might lead others to view an individual as “autistic” (O’Brien, 1981). In fact, efforts at early intervention for children with PDDs have often included strategies for the reduction of stereotypy due to their inverse relationship with social interactions (e.g., Lee & Odom, 1996; Risley, 1968; Singh, Dawson, & Manning,
appropriate play (e.g., Epstein, Doke, Sajwaj, Sorrell, & Rimmer, 1974; Koegel, Firestone, Kramme, & Dunlap, 1974; Thompson & Berkson, 1985), acquisition of discrimination tasks (e.g., Koegel & Covert, 1972; Lovaas et al., 1966), responding to auditory stimuli (e.g., Lovaas, Litrownik, & Mann, 1971), and acquisition of imitative tasks (e.g., Risley, 1968). Researchers have also found that the presence of stereotypy is associated with higher levels of stress for families of children under the age of four with disabilities (Beckman, 1983). Of greater concern is that some stereotypy may be related to the emergence of self-injurious behavior (Berkson & Tupa, 2002; Kennedy, 2002) and other aggressive behavior (Fisher, Lindauer, Alterson, & Thompson, 1998). In addition, some stereotypy such as bruxism or mouthing/licking one’s skin may pose health risks such as dental malocclusion or skin deterioration (Rojahn & Sisson, 1990).

Behavioral Interventions

A number of behavioral interventions (e.g., extinction, time out, environmental enrichment, differential reinforcement) have shown some efficacy in reducing stereotypy. However, not many direct comparison studies have been conducted, and most studies have not reported generalization or maintenance data, with the exception of studies of overcorrection (Schroeder, 1991). In one direct comparison study, Harris and Wolchik (1979) found that four participants responded differently to some of the same treatments. The equivocal results found in this study and in the literature may be a result of not conducting pre-treatment functional assessments. Without a functional assessment, there is no basis for predicting which treatment will be most effective. Many of the following treatments may be more effective when matched to the function of the stereotypy.
When stereotypy is maintained by social reinforcers, treatment may resemble that of other behaviors similarly maintained (e.g., attention-maintained self-injury). However, even when a social function is present, automatic functions are often concurrently present (e.g., Kennedy et al., 2000). Automatically reinforced behaviors are often difficult to treat because the sensory reinforcer cannot be easily withheld or blocked by the therapist (Vollmer, 1994). The most commonly reported interventions for stereotypy can be categorized as antecedent manipulations, differential reinforcement, extinction, punishment, a combination of these, and a unique treatment called reinforcer displacement that comprises elements of both reinforcement and extinction. Below is a summary of interventions from each of these categories.

**Antecedent Manipulations**

This category of interventions includes procedures that alter the individual’s physiological state or the environment to change the discriminative stimuli (S<sup>D</sup>s) or establishing operations (EOs) that evoke stereotypy, as opposed to delivering or withholding consequences following their occurrence. While EOs alter the effectiveness of specific events as reinforcers and evoke behaviors that have been reinforced by those events, S<sup>D</sup>s evoke behaviors that have been differentially reinforced in the presence of those stimuli (Michael, 1993).

One type of antecedent intervention, pharmacotherapy, has sometimes been successful in treating stereotypy. However, “the pharmacotherapy of [stereotypic movement disorder] and of stereotypies remains in its infancy” (Stein & Simeon, 1998, p. 330) and more research is needed on neuroleptics and other drugs prescribed for stereotypy (Rojahn & Sisson, 1990). In addition, pharmacotherapy is often not a
preferred treatment due to side effects (Rojahn & Sisson, 1990) and has been shown to be less efficacious than behavioral treatments (Singh, Landrum, Ellis, & Donatelli, 1993), especially in the treatment of stereotypy (Lennox, Miltenberger, Spengler, & Erfanian, 1988). In a study conducted by Singh, Watson, and Winton (1987), parents chose "behavior modification" techniques over drug therapy when given a choice of four treatments for their children.

Physical exercise, relaxation, and massage have also been used as antecedent treatments for stereotypy (e.g., Escalona, Field, Singer-Strunck, Cullen, & Hartshorn, 2001; Morrissey, Franzini, & Karen, 1992; Powers, Thibadeau, & Rose, 1992). These interventions may work by providing kinesthetic, vestibular, or other types of sensory stimulation that reduces the EO(s) that evoke stereotypy. Although some reductions have been observed, the best clinical application of these techniques is likely in conjunction with other procedures (Rojahn & Sisson, 1990).

Environmental enrichment (EE) likely achieves its effects by reducing the EO that evokes stereotypy maintained by automatic reinforcement. Although some studies have not been based on hypotheses of maintaining variables, EE may be more effective when analyses are conducted which not only identify an automatic reinforcement function, but also identify the specific type of sensory stimuli maintaining the stereotypy (LeBlanc, Patel, & Carr, 2000). For example, Piazza, Adelinis, Hanley, Goh, and Delia (2000) conducted functional analyses of various problem behaviors and identified automatic reinforcement functions for each. The authors then conducted stimulus preference assessments with stimuli matched (i.e., similar) and unmatched (i.e., dissimilar) to hypothesized sensory reinforcers. When presented noncontingently, stimuli matched to
sensory hypotheses resulted in greater reductions in stereotypy than unmatched stimuli. Although sometimes effective, these additional assessments may be impractical to implement due to the time, expertise, resources, or high level of environmental control that are often needed (LeBlanc et al., 2000).

Other antecedent interventions for stereotypy have included behavioral momentum (e.g., Mace & Belfiore, 1990) and curricular revision by using shorter inter-trial intervals (e.g., Dunlap, Dyer, & Koegel, 1983), varied tasks (e.g., Winterling, Dunlap, & O’Neill, 1987), or prompts to complete a task (e.g., Symons & Davis, 1994). These studies show promise for the treatment of stereotypy, but more research is needed.

Differential Reinforcement

Reinforcement-based procedures for treating stereotypy include differential reinforcement of other behavior (DRO; e.g., Harris & Wolchik, 1979; Koegel & Koegel, 1990), incompatible behaviors (DRI; e.g., Paisey, Whitney, & Wainczak, 1993; Weisberg et al., 1973), alternative behaviors (DRA; e.g., Mulhern & Baumeister, 1969), and low rates of behavior (DRL; e.g., Singh et al., 1981). Overall, these procedures have not been effective in eliminating stereotypy when used alone (Rojahn & Sisson, 1990), have not shown as rapid or durable effects as other treatments, and have had best results used in conjunction with punishment-based procedures. As a group, differential reinforcement procedures may not be as effective as one would expect because of the difficulty of identifying reinforcers powerful enough to compete with those maintaining stereotypy and that do not result in rapid satiation (Koegel & Koegel, 1989).
**Punishment**

Although not usually a preferred method of treatment, punishment-based interventions such as programmed aversive stimulation and overcorrection have been the most effective in reducing stereotypy. Aversive stimulation has been delivered contingent upon stereotypy in the form of electric shock (e.g., Baumeister & Forehand, 1972; Lovaas, Schaeffer, & Simmons, 1965), ice (e.g., Drabman, Ross, Lynd, & Cordua, 1978), smelling salts (e.g., Clarke & Thomason, 1983), shaking (e.g., Risley, 1968), water mist (e.g., Bailey, Pokrzywinski, & Bryant, 1983; Reilich, Spooner, & Rose, 1984), immobilization (e.g., Bitgood, Crowe, Suarez, & Peters, 1980; Reid, Tombaugh, & Heuvel, 1981), facial/visual screening (e.g., Jordan, Singh, & Repp, 1989; McGonigle, Duncan, Cordisco, & Barrett, 1982), among others. Interventions based on punishment have been shown to produce rapid and substantial decreases in stereotypy (Koegel & Koegel, 1989). Most procedures have been shown to be highly effective, with electric shock being the most effective (LaGrow & Repp, 1984). Although these techniques do not frequently produce the negative side effects often attributed to them, aversive interventions are perhaps more acceptable to caregivers as a last resort, and are most effective when used according to very specific protocols and by very skilled therapists (Smith, 1990).

Overcorrection, first used by Foxx and Azrin (1973), involves requiring an individual to repeatedly perform an appropriate behavior immediately following an inappropriate behavior. Many studies have demonstrated its effectiveness in reducing stereotypy, with some also noting positive side effects such as increases in appropriate play, smiling, and learning. However, some researchers have also reported negative side
effects of increased problem behaviors and minimal transfer to other environments (O’Brien, 1991). In addition, overcorrection may not be appropriate for use in all settings because it may be difficult for staff to implement (Rojahn & Sisson, 1990) and has high potential for misuse (Koegel & Koegel, 1989).

Although time-out has occasionally been effective in reducing stereotypy (e.g., Pendergrass, 1972), in most cases it has been found to have no effect or increase stereotypy (e.g., Clarke & Thomason, 1984; Harris & Wolchik, 1979; Koegel & Koegel, 1989). Time-out, which is typically defined as the removal of access to reinforcement for a brief period of time contingent upon a target behavior, will likely be ineffective unless matched to behavioral function. In fact, it may come to function as a reinforcer for the behavior if it provides the individual with more time to engage in stereotypy, or if stereotypy is maintained by escape from demands.

**Extinction**

Because extinction-based procedures work by discontinuing the response-reinforcer relationship, their implementation without knowledge of behavioral function is difficult, if not impossible. Stereotypy maintained by social reinforcers may be extinguished by withholding attention, tangibles, or through the continued presentation of tasks. However, utilizing these procedures without first determining the function of stereotypy may produce no effect or actually increase stereotypic behaviors. Sensory extinction may be much more difficult to implement because specific reinforcing sensory characteristics must be identified. For example, Aiken and Salzberg (1984) blocked auditory stimulation from stereotypic vocalizations by playing white noise. Other studies have blocked proprioceptive (e.g., Rincover, Cook, Peoples, & Packard, 1979), auditory
(e.g., Rincover, 1978), and visual/kinesthetic (e.g., Maag, Wolchik, Rutherford, & Parks, 1986) stimulation to reduce stereotypy. Although demonstrated to be efficacious, sensory extinction may not be practical because of the highly controlled environment necessary for its assessment/implementation and the difficulty of eliminating all of the sources of automatic reinforcement often inherent in stereotypy (Rojahn & Sisson, 1990).

Reinforcer Displacement

Neisworth, Hunt, Gallop, and Madle (1985) introduced the term reinforcer displacement (RD), based on the assumption that “it might be possible to shift the control of a behavior from some historic and inaccessible partial schedule to a clinically imposed continuous one, thus providing optimal circumstances for weakening the behavior through extinction” (p. 104). This is a description of the interpolated reinforcement effect (IRE), which, although supported by some studies with humans (e.g., Higbee, Carr, & Patel, 2002), has been shown in more recent research to be only minimally effective as a reductive procedure (Carr, 2002). Based on the partial reinforcement extinction effect, IRE procedures involve changing the schedule from intermittent reinforcement to continuous reinforcement (CRF) prior to extinction.

However, RD differs from the IRE procedurally, and likely, mechanistically. The RD procedure involves delivering a new type of stimulus contingent on behavior that is already maintained by some other controlling stimulus (e.g., automatic reinforcer) prior to removal of the new stimulus. A change in schedule of reinforcement from intermittent to continuous is not part of this intervention. In fact, in the case of behaviors maintained by automatic reinforcement, the already maintaining reinforcer is likely not being “delivered” on an intermittent schedule of reinforcement, but rather a continuous and
possibly conjugate one. The potential applied significance of the RD procedure is that it could be used to shift the control of a behavior from a difficult-to-manipulate stimulus to an easily manipulated stimulus, such as a tangible item, allowing a clinician to more easily implement extinction. For this reason, RD may be a particularly useful treatment for stereotypy.

Because RD might result in a decrease in behavior following removal of the new stimulus, it resembles what some researchers have referred to as the “undermining of intrinsic motivation” (Deci, Koestner, & Ryan, 1999). Some authors maintain that when extrinsic reinforcers are delivered contingently upon behaviors that are intrinsically motivated, intrinsic motivation is “corrupted,” resulting in a decrease in those behaviors when the reinforcer is withdrawn. This hypothesis has been the source of much controversy for the past 30 years (Deci et al., 1999). However, a meta-analysis of approximately 100 studies by Cameron and Pierce (1994) found that these effects only occurred when expected tangible rewards were delivered merely for engaging in the behavior rather than for high-level performance. However, it cannot yet be ascertained if this process is the same as RD, as the mechanisms by which each operate have not been determined.

Foxx and McMorrow (1983) conducted the first study using the RD technique with two adults diagnosed with profound mental retardation. In a series of 3 experiments, reversal designs were used to evaluate the effects of RD, consisting of baseline, a period of continuous tangible delivery (“reinforcement”; 25 sessions), followed by what the authors termed “extinction.” The use of the terms reinforcement and extinction in many studies on RD implies that the new tangible stimuli did, in fact, come to maintain the
target behavior. In a later study, the procedural terms superimposition and removal of superimposition were used to maintain technical integrity and avoid implications of controlling stimuli. These terms will, therefore, be used throughout this article.

Prior to the experimental phase in the Foxx and McMorrow (1983) study, stimuli were demonstrated to be reinforcers when used to strengthen other behaviors. During baseline, RJ engaged in stereotypic gulping at a mean of 9 responses per min. LL’s baseline rate of zipper play occurred at a mean of about 5.25 responses per min. During superimposition, RJ’s rate of gulping remained the same, and then decreased dramatically to a mean of 2.5 responses per min during the final removal of superimposition phase. Although the data suggest a possible upward trend in the final removal sessions, there are not enough data to form conclusions concerning durability of the effect. Interestingly, LL’s rate of zipper play did not increase during superimposition, possibly due to time spent consuming reinforcers or because of a ceiling effect. During removal, stereotypy increased abruptly in a “burst”, followed by a return to baseline levels. In a third experiment, LL was exposed to both continuous and VR5 superimposed tangible delivery phases prior to tangible removal. Continuous delivery resulted in the same pattern of behavior observed with LL in the previous experiment (i.e., “burst” and return to baseline levels). Although stereotypy reduced to 1 response per min during one session of removal following VR5, it immediately returned to baseline levels in the next session.

Although the superimposition procedure produced a decrease in RJ’s stereotypy, it is not clear why LL’s behavior followed a different pattern. It is not possible to draw conclusions concerning the durability of the effect of the superimposition procedure with RJ because of the brevity of this phase. In addition, because a functional analysis was not
conducted, it is not clear if social functions were present and played a role in these results.

Neisworth et al. (1985) evaluated the superimposition technique with two 19-year-old males, each diagnosed with severe mental retardation. Both individuals displayed little vocal language, functioned at a level low enough to be considered untestable, and displayed high rates of stereotypy. Matthew engaged in hand flapping against objects at a mean of approximately 70% of intervals. Stephen’s finger flicking occurred at a mean of approximately 52% of intervals during baseline. Prior to the experiment, tangible stimuli were identified as reinforcers by training new responses. An A-B-A design was used to evaluate the effects of the superimposition procedure, consisting of baseline, continuous superimposed tangible delivery (25 and 28 sessions, respectively), and removal of superimposition. Sufficient and consistent item delivery was ensured by delivering 25 tangibles during each session of the superimposition phase. As Foxx and McMorrow (1983) observed, the rate of stereotypy did not increase much during this phase. For both participants, however, responding became much less variable.

Substantial decreases in stereotypy were observed for both Matthew and Stephen during extinction. Matthew’s handflapping decreased by about 64% from mean baseline levels, and continued to decrease to a mean of 3% of intervals at a two-week follow-up. Stephen’s behavior, however, followed a different pattern. Although finger flicking decreased to a mean of 30% of intervals, it had been at near-zero levels for 16 sessions of removal, and then returned to baseline levels and remained there two weeks later.

There are several considerations in evaluating the results of the Neisworth et al. (1985) investigation. Although this procedure produced substantial decreases in
stereotypy, it is unclear why this effect was only durable for one of these participants. Because a functional analysis was not conducted, it cannot be determined what role behavioral function played in these discrepant results. Anecdotally, the authors noted that when Matthew emitted stereotypy, he often then looked at the experimenter and put out his hand. This suggests that food may have been controlling hand flapping (during superimposition). However, it also raises a question about whether the food receptacle was present in baseline and removal conditions—if not, the presence of the food receptacle during the superimposition phase may have come to signal the availability of food and thus function as an $S^D$ for hand flapping. In addition, no information is provided about collateral behaviors/stereotypy or about generalization to other environments.

The Neisworth et al. investigation, although published in 1985, was replicated and extended by one of Neisworth’s students as a dissertation during the previous year (Hunt, 1984). Utilizing a multiple-baseline design across four participants, Hunt evaluated the superimposition procedure alone, in conjunction with DRI, with different superimposition phase lengths, in different environments, and in terms of changes in collateral behaviors. Participants ranged from 10 to 21 years in age, were diagnosed with moderate to profound mental retardation, exhibited relatively high rates of stereotypy in multiple settings, and had no or limited mand repertoires. Targeted stereotypic behaviors included repetitive hand movements, foot stomping, inappropriate vocalizations, and mouthing/flapping. These behaviors occurred at mean baseline levels ranging from approximately 27% to 48% of intervals. Prior to treatment, stimuli were verified as reinforcers when used to strengthen arbitrary low-rate behaviors. The same experimental
design used in Neisworth et al. (1985) was employed: baseline, superimposition, and superimposition removal. However, the superimposition phase of the procedure varied at 25, 35, 47, or 64 sessions in length. Again, delivery of tangibles on a continuous schedule contingent upon stereotypy resulted in a less variable pattern of responding for all participants. In addition, although stereotypy was moderately elevated in the superimposition condition, almost all sessions were still within baseline levels. Unlike the Neisworth et al. investigation, Hunt found that stereotypy only decreased during extinction for one participant, Trish. The length of superimposition phase did not seem relevant, as Trish was exposed to 47 sessions of superimposition, and another participant was exposed to 64 sessions. Duration measures were not reported for Trish because DRI was implemented immediately following the decrease. DRI was also implemented for another participant and resulted in a substantial decrease in stereotypy. Interestingly, other motor stereotypy decreased below baseline levels for all participants in the final phases, but vocalizations increased or did not change. In addition, self-injurious behavior present during baseline was eliminated in final phases. Probes in the classroom environment showed that target stereotypy did increase in this environment, an important consideration in evaluating the superimposition procedure. Also important to note is the lack of any problem behavior during removal procedures and the ease with which the authors note this procedure was implemented.

Although these results are very different from those reported by Neisworth et al. (1985), several issues are worth noting. First, this dissertation was never published, and therefore has not had the opportunity to be evaluated through the peer-review process. Second, whereas Neisworth et al. ensured consistency of number of tangibles delivered
per session across participants, Hunt (1984) did not. Tangibles during superimposition were delivered after each occurrence of stereotypy, a rate that varied each session for each participant. This may have resulted in different amounts of exposure to superimposition than was planned. In reference to stereotypy in the classroom environment, it is also not clear if the experimenter present during experimental sessions was also present during these probes. If so, the experimenter’s presence could have been discriminative for stereotypy. Third, interobserver agreement measures were only collected during 7% of sessions for three participants, and 12% for one participant. Finally, a functional analysis was not conducted, which could possibly have helped explain the different results between Trish and the other participants.

Schmid (1986) also extended the findings of Neisworth et al. (1985). An important distinction was that behaviors in the context of academic demands were targets rather than stereotypy. These problem behaviors, such as turning away from tasks as well as other “off-task” behaviors, may have been escape-maintained, but a functional analysis was not conducted to confirm this possibility. This would make the treatment procedure more similar to the IRE procedure because these behaviors were perhaps intermittently reinforced in the natural environment and then subsequently exposed to continuous tangible delivery. In addition, if these behaviors were maintained by escape from an aversive teaching interaction, presenting preferred items would perhaps alter the EO that previously evoked the escape responses. Participants were 3 males and 3 females, ranging in age from 10 to 16 years and diagnosed with mild to severe mental retardation. Information on language abilities was not reported, but participants were able to perform training tasks in their rehabilitation programs. The procedure was implemented in the
same manner as in previous studies (i.e., baseline, superimposition, removal), but
superimposition sessions were considerably briefer, ranging from 1 to 8 sessions.
Tangibles and praise used in the superimposition sessions were not demonstrated to be
functional reinforcers prior to the experiment, but during superimposition responding
moderately increased. During the removal phase, problem behaviors decreased for all
participants; however, the durability of these reductions is unknown because the removal
phase was brief, lasting only 3 to 7 sessions across 1 to 2 days.

Clearly, Schmid (1986) demonstrated that the procedure used was effective in
reducing problem behavior, even with very brief exposure to superimposition with
tangibles. However, effects on other settings or collateral behaviors were not evaluated.
In addition, it is not clear if behavioral function was targeted by the intervention, because
the variables originally controlling behaviors were not identified.

Foxx, McMorrow, Fenlon, and Bittle (1986) evaluated tangible superimposition
as part of an intervention to reduce the genital stimulation of Zeke, a boy diagnosed with
severe retardation. No language or skill level information was reported for Zeke. In a
multiple-baseline design across school settings, genital stimulation was first reduced with
a graduated DRO with the “reinforcer” being access to stereotypical rice play (i.e.,
repeated dropping and picking up uncooked rice). During baseline, genital stimulation
occurred at a mean of approximately 13% of intervals in the morning class and 17% of
intervals in the afternoon class. During the DRO procedure, each instance of rice
dropping was followed by the delivery of chocolate cereal (i.e., continuous tangible
superimposition). This continued for 22 sessions both in the morning and the afternoon.
As the DRO time increased, time of access to rice play also increased until genital
stimulation was reduced to near zero rates in both settings, rice play was offered on a DRO 45 min for a duration of 7.5 min, and occurred during approximately 5% of intervals. Next, the effects of superimposition on rice play were evaluated when the chocolate cereal was withheld (i.e., removal of superimposition). The resulting data showed no effect of superimposition removal on rice play or genital stimulation.

One point worth noting in evaluating the failure of the superimposition procedure in the Foxx et al. (1986) study is that the chocolate cereal was not delivered directly to Zeke. Rather, pieces of cereal were delivered into a cup, to which Zeke was given access at the end of the sessions. It is unclear whether chocolate cereal was a preferred item for Zeke. In addition, it is unclear whether the sight of the cereal being dropped into the cup functioned as a conditioned reinforcer for rice play or Zeke was able to engage in rule-governed behavior that would bring rice play under the control of cereal delivery. Furthermore, the function of rice play was not assessed in this study, so it is unknown if the failure of the superimposition procedure was related to behavioral function.

Wylie and Grossman (1988) evaluated the RD procedure by replicating the Neisworth et al. (1985) study with eight rats. Interestingly, although this was a basic study with nonhumans, it was published in *Journal of Applied Behavior Analysis* because of its potential applied relevance for the problem behavior of humans. An A-B-A design was used to evaluate the effects of the superimposition procedure, consisting of baseline, continuous tangible superimposition (15 sessions), and removal of superimposition. Lever pressing was shaped operantly and maintained by the delivery of food pellets. During baseline, 4 of the rats were exposed to 25 sessions of a variable-ratio (VR) schedule of reinforcement and 4 of the rats were exposed to a variable-interval (VI)
schedule of reinforcement. In the VR group, baseline responding was established under a VR-25 schedule for two of the rats and under a VR-75 schedule for the other two rats. Similarly, in the VI group, baseline responding was established under a VI 25-s schedule for two of the rats and under a VI 75-s schedule for the other two rats.

Although other studies have evaluated the superimposition procedure in the context of the delivery of a different type of tangible, Wylie and Grossman (1988) used only food throughout all phases of the study, but superimposed a different schedule of food delivery (i.e., continuous) on top of an already existing schedule of reinforcement. Although similar to the IRE procedure, this differs in that the schedule is not being changed from an intermittent one to a continuous one. Rather, a continuous schedule is being added or superimposed on to an intermittent schedule. In other words, during the continuous superimposition phase, both schedules were simultaneously in place. Following 15 sessions of continuous superimposition, the CRF schedule was removed in a reversal back to intermittent reinforcement baseline conditions.

During continuous superimposition, lever pressing was substantially decreased and variability was reduced. It is likely that this was because the rats did not continue to lever press while consuming the food. When the continuous schedule was removed, responding quickly returned to baseline levels. These patterns were observed across all subjects, and did not differ as a function of baseline schedule of reinforcement. Although unlikely, it is possible that the discrepant results obtained in this study may be due to the fact that the superimposition effect is not as readily observed in nonhuman populations. It is also possible that the type of reinforcer superimposed must be of a different type, not just delivered on a different schedule, for the effect to be observed. Another
consideration with these results is that when continuous superimposition was removed, intermittent reinforcement was still delivered for lever pressing. In the case of behaviors maintained by automatic, positive reinforcement in the form of attention or negative reinforcement, the superimposition procedure may work by changing some aspect of the original contingency or the quality of the reinforcement. The intermittent reinforcement delivered during the removal phase would not accurately represent this situation. Due to these issues, the authors called for more research on this phenomenon, with a focus on determination of the type and schedule of reinforcement maintaining the target response.

Although other treatments such as punishment, sensory extinction, and EE have demonstrated some effectiveness in reducing stereotypy, research indicates that their utility may be limited due to practical and ethical concerns. In addition, these interventions have not been found to be consistently effective for all individuals (LeBlanc et al., 2000; Rojahn & Sisson, 1990; Smith, 1990). The limited data from the studies evaluating the superimposition procedure suggest that it may also be a viable option for the treatment of stereotypy. Because it is not a function-based treatment, the resources needed to conduct a functional analysis would not be required, as is necessary with sensory extinction and some EE procedures. Further, it could be used when a problem behavior’s maintaining reinforcer cannot be identified or manipulated (as in some cases of automatic reinforcement). However, given that only five published studies have evaluated the intervention, further research is needed to evaluate the efficacy and durability, as well as the role of behavioral function, in the superimposition procedure. Because the procedure involves the delivery of tangibles contingent upon problem behavior with intent to produce a reinforcement effect, ultimately clinical use of this
procedure would depend on the development of practical methods that result in a robust
and reliable production of this phenomenon.

The purpose of the current study was to replicate and extend previous research on
the superimposition procedure by evaluating its efficacy and durability for high-level
stereotypy. Extensions of previous studies will include implementation of this procedure
with a different but clinically relevant population, as well as more stringent experimental
procedures and interobserver agreement measures. In addition, the functions of
stereotypic behaviors were experimentally determined prior to treatment to allow for a
closer examination of potential discrepant results.

METHOD

Participants

Three children diagnosed with autism participated in the study. This number of
participants was selected due to the requirements of the single-case design that was
employed (Barlow & Hersen, 1984). Although the superimposition procedure has been
used to treat older individuals with different diagnoses (e.g., adolescents with mental
retardation), the proposed age group and diagnosis are appropriate for this study because
these are characteristics of children often targeted for intensive, early intervention
programs. In addition, stereotypy is one of the diagnostic criteria for autism. As
previously mentioned, high rates of stereotypy have been shown to be inversely related to
several target behaviors of common early intervention programs, including play, social
interactions, discriminations, among others. Only children with high-rate stereotypy that
interfered with adaptive functioning were included in the study. Inclusion criteria were
based on an informal review of the literature, which indicates that researchers have
generally evaluated treatments for individuals whose baseline means of stereotypy were between 40% and 100% of intervals.

Rose was 6-years old and engaged in high rates of scratching objects and mouthing. In addition to autism, Rose was also diagnosed with a seizure disorder. Rose did not display vocal-verbal behavior, with the exception of saying “eh” when desired objects were out of reach. Her adaptive behavior score on the Vineland Adaptive Behavior Scales, Interview Edition, Survey Form (VABS; Sparrow, Balla, & Cicchetti, 1984) was 29, indicating a severe deficit. Her age-equivalent scores across domains ranged from 11.0 months to 1.5 years. Libby was 6-years old and engaged in high rates of scratching objects, mouthing, and pressing her body against surfaces. Although Libby displayed frequent and varied vocalizations, she did not display any vocal-verbal behavior. Her adaptive behavior score on the VABS was 31, indicating a severe deficit. Her age-equivalent scores across domains ranged from 6 months to 1.2 years. Doug was 5-years old and engaged in hand flapping and inappropriate vocalizations. According to the Behavioral Language Assessment (BLA; Sundberg & Partington, 1998) conducted with his caregiver, Doug consistently displayed a verbal repertoire that consisted of more than 10 vocal mands, echoics with many sounds and words, 16 to 50 different tacts, and between 10 and 20 simple intraverbals. His adaptive behavior score on the VABS was 32, indicating a severe deficit. His age-equivalent scores across domains ranged from 11 months to 2.5 years.

**Dependent Variables and Data Collection**

Sessions were conducted in a quiet area of each child’s home or school 4-5 days a week and were 15 min in duration. Large furniture remained in the room during sessions,
but smaller items were removed or covered. During baseline, superimposed tangible delivery, and removal of superimposition sessions, one graduate experimenter and one undergraduate assistant were present along with a timer, a container for snacks, and a video camera. During the functional analysis, toys and colored poster board were also present. Undergraduate research assistants videotaped and scored all sessions.

Rose and Libby’s target stereotypy was *scratching*, which was defined as movement of finger tip(s) and/or fingernail(s) across an object without using the same fingers to grasp the item. For Rose this also included her toes. Doug’s target stereotypy was *hand flapping*, defined as repetitive movement of hand(s) other than to manipulate an object. Data on other problem behaviors were also recorded during all phases of the study. Rose exhibited pinching, Libby exhibited whining, and Doug did not display any problem behavior throughout the study. During the EE treatment evaluation, data were also collected on toy engagement. Generally, this was defined as touching the toy without scratching it; however, specific definitions for toy engagement were used for each toy. For example, toy engagement was scored for the bead rattle if Libby held it for longer than 2 s, shook it, or tapped it on something without scratching it. For the functional analysis and treatment evaluation, behaviors were recorded from videotape using a noncontinuous partial-interval recording system (10 s observation; 5 s recording).

*Informant and Direct-Observation Assessment*

An initial screening interview was conducted to assess age, diagnosis, frequency of stereotypy, and extent to which stereotypy interfered with adaptive functioning. This interview also included questions about environmental antecedents and consequences of stereotypy based on the Functional Assessment Interview (FAI; O’Neill et al., 1997).
Several 15-min observations were conducted to determine that at least one topography of stereotypy occurred at a mean rate between 40% and 80% of 10-s intervals. An upper limit of 80% was set in order to later demonstrate a reinforcement effect during treatment. If a child met the criteria listed above, subsequent observations were conducted to obtain more specific information about the topographies of stereotypy and collateral behaviors, as well as the events that preceded and followed stereotypy, so that a functional analysis could be conducted. The target for each child was the highest frequency stereotypic behavior that was compatible with consumption of the food item that would be delivered during superimposition sessions (e.g., did not involve the mouth).

**Skills Assessment**

*Adaptive behavior assessment.* The VABS (Sparrow et al., 1984) is designed to assess the adaptive behavior of children with and without disabilities. This semi-structured interview assesses functional performance in areas of language skills, social skills, motor skills, and everyday living, as well as problem behavior. In the current study, this assessment was conducted to provide a broad and systematic description of adaptive-behavior characteristics of each participant.

*Language assessment.* The communication scale of the VABS and the BLA were used to assess the communication skills of each participant so that discrepant results could be analyzed in terms of language ability. Based on Skinner’s (1957) analysis of verbal behavior, the BLA is a 12-item interview that assesses language functionally rather than topographically, providing perhaps a more complete representation of a range of different types of verbal behavior.
**Functional Analysis**

Functional assessments have not been conducted in previous studies on RD, as knowledge of function is not necessary to implement this treatment. However, because social, automatic, or multiple functions of stereotypy may produce differential treatment outcomes, functional assessments were conducted for each participant in the present study. This would also provide information on reinforcer type, which Wylie and Grossman (1988) indicated might be critical in producing this phenomenon.

Based on information obtained from the informant and direct-observation assessments, a functional analysis was used to experimentally identify the function of target stereotypy. The functional analysis was evaluated using a multielement design using procedures similar to those described by Iwata et al. (1982/1994). However, each participant’s conditions were created based on individual information obtained from each participant’s informant assessment. All functional analyses included an **attention** condition to test for sensitivity to positive reinforcement in the form of attention, a **no-interaction** condition to test for sensitivity to automatic reinforcement, and a **control** condition consisting of no demands, noncontingent attention, and free access to tangibles. Rose and Libby’s functional analyses also included a **demand** condition to test for sensitivity to reinforcement in the form of escape from demands. Rose’s functional analysis also included a **tangibles** condition to assess a possible tangible reinforcement function. These conditions were included because prior observation and informant assessment indicated that these reinforcers were sometimes delivered contingent on stereotypy in the natural environment.
A multiple-stimulus (without replacement) preference assessment (MSWO; DeLeon & Iwata, 1996) was used to identify preferred toys for the attention, tangible, and control conditions. Toys were selected which were not associated with high rates of the target stereotypy. During this assessment, participants were asked to choose from a linear array of eight toys. After the participant had 10 s of access to the selected item, the toy was removed from view. The remaining items were then rearranged and a new instruction was issued. This procedure was repeated until all toys were chosen or the participant did not indicate preference for any toys. The assessment was repeated two more times, yielding a total of three arrays (Carr, Nicolson, & Higbee, 2000).

All functional analysis sessions were 15 min in duration (Wallace & Iwata, 1999; 10 min for observation, 5 min for recording), with conditions presented in a quasi-random sequence until responding across conditions reached visual stability. Different colored poster boards were present in the room in a salient location to help facilitate discrimination between conditions (Conners et al., 2000). Upon entering the room, the experimenter announced the session number and tapped the poster board to draw the participant’s attention to it.

During the no-interaction condition, experimenters did not interact or make eye contact with the participant except to ensure his or her safety. No additional materials were present in the room. During the attention condition, the participant was provided with three toys ranked as 4, 5, and 6 in the preference assessment. For Rose, this included 2 different books and a cow noisemaker; for Libby, this included blocks, a book, and a small plastic spoon; for Doug, this included a toy car, a book, and a Magnadoodle®. The experimenter provided eye contact, close proximity, and mild social disapproval
(e.g., “Hands down”, “What’s wrong? Why are you doing that?”) contingent on target stereotypy. The topography of the social disapproval was individualized for each participant based on parental report and prior direct observation of parent behavior. During the control condition, the experimenter and the participant played with toys ranked as 1, 2, and 3 in the preference assessment. For Rose, this included a book, a fabric block, and a shaker; for Libby this included a tambourine, plastic figurines, and a puzzle; for Doug, this included 2 books and Mr. Potatohead®. During this time, noncontingent attention was delivered either on a fixed time 30-s (Rose and Libby) or informal schedule (Doug). No demands were made of the participant and no differential consequences were provided for stereotypy. During Rose and Libby’s demand condition, the experimenter delivered verbal instructions and least-to-most prompts for imitation tasks (e.g., “Touch head” with verbal, then imitative, then physical prompting). The participant was given a 20- to 30-s break contingent on occurrence of target stereotypy. If stereotypy occurred during the break, the period was reset. During Rose’s tangibles condition, she was given access to a ribbon for 30 s contingent on the occurrence of target stereotypy.

In the event that responding was undifferentiated across sessions, several extended no-interaction sessions were conducted to confirm that the stereotypy was not maintained by social reinforcement. This is necessary because undifferentiated patterns of responding may also result from multiple-treatment interference or intermittent social reinforcement (LeBlanc et al., 2000).
Interobserver Agreement

Interobserver agreement (IOA) on occurrence of stereotypy was assessed using the overall (i.e., point-by-point) agreement method (# of agreements / [# of agreements + disagreements] x 100%). An agreement was defined as both observers recording the target stereotypy, problem behavior, or appropriate toy play during an observation interval. IOA was assessed for 58% of Rose's functional analysis sessions and 52% of her treatment evaluation sessions. Mean IOA for scratching was 92.4% (range, 84%-98%) and 95.4% (range, 77%-100%) for the functional analysis and treatment evaluation, respectively. IOA was assessed for 32% of Libby's functional analysis sessions and 30% of her treatment sessions. Mean IOA for scratching was 96% (range, 87%-100%) and 90% (range, 83%-100%) for the functional analysis and treatment evaluation, respectively. Mean IOA for hand flapping was 97.5% (range, 93%-100%) for Doug's functional analysis and was assessed during 44% of sessions.

Treatment Evaluation

The treatment evaluation could not be conducted with Doug because he consistently refused to come with the experimenters to sessions after the functional analysis.

Preference assessment. Caregivers were asked to complete a questionnaire based on the Reinforcer Assessment for Individuals with Severe Disabilities (RAISD; Fisher, Piazza, Bowman, & Amari, 1996). This modified version of the RAISD asks the caregiver to list and rank 10 foods that his or her child seems to prefer. Eight of these items were used in an MSWO assessment, conducted as previously described. Highly preferred foods were delivered contingent upon stereotypy during the superimposition
phase. Because Libby did not respond differentially during the MSWO preference assessment, paired-stimulus (Fisher et al., 1992) and single-stimulus (Pace, Ivancic, Edwards, Iwata, & Page, 1985) preference assessments were subsequently conducted. However, Libby did not display preference for any one food, as she responded by reaching for the closest food during the paired stimulus assessment, and consumed all foods during the single-stimulus assessment.

For Rose, a brief food preference assessment was conducted prior to each session in which the experimenter presented her with a linear array of the five highest preference foods from the MSWO assessment (i.e., raisins, Cheerios®, Cheetos®, m&ms®, and cheese goldfish crackers). After giving Rose one small piece of the food she chose, the experimenter placed this container of food in a bag to be used as putative reinforcers during the superimposition phase. Although no consequences were delivered during baseline or superimposition removal, the bag was present during all sessions to ensure that its absence did not function as an S-delta for reinforcement of stereotypy. Only one food, popcorn, was used with Libby because of the previously mentioned preference assessment results and because other reportedly preferred foods were difficult to deliver in this manner (e.g., applesauce).

**Experimental design and measurement.** The effects of the superimposition procedure on stereotypy were evaluated in the context of a nonconcurrent multiple-baseline design across participants (Watson & Workman, 1981). Each participant was exposed to baseline, continuous superimposition, and superimposition removal conditions in an A-B-A format.
Data on occurrence of the target stereotypy, collateral stereotypy, and problem behavior were collected using noncontinuous partial-interval recording. Sessions were 15 min in duration (10 min for observation, 5 min for recording). During each session, an experimenter remained near the participant with the food container, without talking or making eye contact with the participant to ensure that behaviors were not unintentionally reinforced. Problem behavior was ignored unless there was danger posed to the child or therapist. Sessions were conducted until visual stability was achieved.

**Baseline.** Baseline sessions were conducted as described above, with no contingencies in place for stereotypy. Sessions were conducted until visual stability was demonstrated and enough sessions were achieved to complete that tier of the nonconcurrent multiple baseline design.

**Continuous tangible superimposition.** During this condition, the food item chosen during the pre-session preference assessment was placed directly into the child’s mouth contingent upon target stereotypy. If the child continued to engage in that behavior while consuming and swallowing the edible, another piece would be delivered after swallowing. This phase continued for at least 25 sessions and until visual stability was observed. This number of sessions has been used in previous studies on the RD procedure (e.g., Foxx & Mc Morrow, 1983; Neisworth et al., 1985).

**Removal of superimposition.** This condition was identical to baseline. Sessions were conducted until visual stability was achieved to assess the durability of the intervention.
Environmental Enrichment

If the superimposition procedure proved ineffective in reducing stereotypy for a participant, the plan was to implement a function-based treatment. Because nonsocial functions were identified for all participants, EE was implemented as treatment for both Libby and Rose's scratching. As previously mentioned, EE is a reductive antecedent intervention in which preferred items are provided noncontingently. EE has been shown to be effective in reducing the stereotypy of some children (e.g., Goh et al., 1995; Piazza et al., 2000).

Preference assessments. Free-operant preference assessments were conducted using toys that matched the hypothesized sensory function of scratching for each participant. For both Rose and Libby, auditory and tactile stimulation were hypothesized to be involved in the maintenance of their stereotypy. Ten toys were evaluated for Rose and 12 toys were evaluated for Libby. The purpose of the free-operant preference assessment was to identify toys associated with the lowest levels of scratching and moderate to high levels of toy engagement (Roane, Vollmer; Ringdahl, & Marcus, 1998). Prior to each session, each participant was provided with 10 s of access to the toy being evaluated during that session. The room was arranged as in baseline with the exception that one toy was placed in the center of the room. Toys were presented 3 times each in a random order. Sessions were 3 min in duration. Data on occurrence of scratching and toy engagement were collected using noncontinuous partial interval recording (2 min for observation, 1 min for recording).

IOA was assessed for 20% of Rose's preference assessment sessions. Mean IOA was 97% (range, 92%-100%) and 100% for scratching and toy engagement, respectively.
IOA was assessed for 31% of Libby's preference assessment sessions. Mean IOA was 94% (range, 75%-100%) and 100% for the scratching and toy engagement, respectively.

**EE treatment evaluation.** During this condition, free access to toys was provided during extended (15 min) sessions to evaluate effects on scratching. The room was arranged as in baseline, with the three toys that resulted in the lowest occurrence of scratching placed around the room. DeLeon, Anders, Rodriguez-Catter, and Neidert (2000) found that this arrangement produced similar decreases in problem behavior and higher levels of toy play when compared with rotating each toy individually. Upon entering the room, the experimenter prompted the child to touch one of the toys. For the remainder of the session, procedures were identical to those used in baseline.

**Independent Variable Integrity**

Treatment integrity was assessed using a noncontinuous partial-interval system. Following each 10-s interval, a plus or minus was scored for correct or incorrect treatment delivery, and summarized as percentage of correct intervals per session. IOA on the delivery of independent variables was calculated using the overall (i.e., point-by-point) agreement method (\# of agreements / [\# of agreements + disagreements] x 100. Interobserver agreement (IOA) on the number of reinforcers delivered was assessed using the total (i.e., frequency ratio) agreement method (low \# / high \# x 100%). An agreement was defined as both observers recording the delivery of a reinforcer.

**Functional analysis.** Data were collected on experimenter delivery of attention, tangibles, or task removal contingent upon each instance of target stereotypy, as well as the absence of these contingences during control and no-interaction conditions. Rose's mean treatment integrity score was 99.8% (range, 95%-100%) and was assessed during
96% of functional analysis sessions. Mean IOA was 100% and was assessed during 57% of these sessions. Libby’s mean treatment integrity score was 99.7% (range, 97%-100%) and was assessed during 92% of functional analysis sessions. Mean IOA was 99% and was assessed during 35% of these sessions. Doug’s mean treatment integrity score was 100% and was assessed during 78% of functional analysis sessions. Mean IOA was 100% and was assessed during 43% of these sessions.

Treatment evaluation. Data were collected on experimenter delivery of tangibles contingent upon each instance of target stereotypy during superimposition sessions, and no delivery of tangibles or attention during baseline, removal, and EE sessions. Rose’s mean treatment integrity score was 99% (range, 98%-100%) and was assessed during 92% of treatment evaluation sessions. Mean IOA was 99.9% and was assessed during 56% of these sessions. Libby’s mean treatment integrity score was 99% (range, 92%-100%) and was assessed during 91% of treatment evaluation sessions. Mean IOA was 99% and was assessed during 31% of these sessions. The number of edible items delivered per session was recorded during 100% of sessions. The mean IOA score on food delivery for Rose was 98% and was assessed during 26% of these sessions. The mean IOA score on food delivery for Libby was 97% and was assessed during 25% of these sessions.

Free-operant preference assessment. Data were collected on experimenter prompts to play with a toy at the beginning of the session and not interacting with the participant at other times during the session except to keep her safe and retrieve a toy if it rolled under the bed. Mean treatment integrity scores for both Rose and Libby were 100% and were assessed during all of the EE preference assessment sessions. For Rose,
IOA was assessed during 27% of these sessions ($M = 100\%$). For Libby, IOA was assessed during 31% of these sessions ($M = 100\%$).

RESULTS

Functional Analysis

The functional analysis graphs for each participant are depicted in Figure 1. As seen in the top panel, Rose's responding was differentially high in the no-interaction condition, low in the control condition, and undifferentiated in the escape, attention, and tangible conditions, ranging from 2\% to 85\% of intervals. Rose's scratching persisted across the 4 consecutive no interaction-condition sessions that were conducted at the end of the analysis. The means for each of Rose's test conditions were as follows: no interaction, 79.4\% ($SD = 4.9$), attention, 38.8\% ($SD = 8.9$), control, 18.3\% ($SD = 17$), demand, 45.5\% ($SD = 14$), and tangible, 37\% ($SD = 17$). As seen in the middle panel, Libby's responding was undifferentiated across conditions during the first 20 sessions, ranging from 3\% to 68\% of intervals. Libby's scratching persisted across the 5 consecutive no-interaction condition sessions that were conducted at the end of the analysis. The means for each of Libby's test conditions were as follows: no interaction, 51.5\% ($SD = 21.2$), attention, 40.8\% ($SD = 19.3$), control, 30.6\% ($SD = 27$), and demand, 45.4\% ($SD = 20.4$). As seen in the bottom panel, Doug consistently exhibited more hand flapping in the no-interaction condition ($M = 49.3\%, SD = 27$) compared to the control and attention conditions ($M = 3.5\%, SD = 3.5, M = 1.3\%, SD = 1.2$, respectively).

The differentially high responding in the no-interaction condition (Rose and Doug) and undifferentiated functional analysis (Libby) suggest that the participants'
stereotypic behaviors were all maintained independent of social consequences.

Stereotypy persistence across multiple consecutive no-interaction sessions (in the absence of social contingencies and alternating test conditions) rules out, to some degree, the possibility that the undifferentiated responding was a result of multiple-treatment interference associated with the multielement design (Vollmer, Marcus, Ringdahl, & Roane, 1995).
Figure 1. Functional Analysis Outcomes for Each Participant.
Treatment Evaluation

Superimposition. The treatment evaluation graphs for Rose and Libby are depicted in Figure 2. As seen in the top panel, Rose’s scratching was variable during baseline and occurred at a mean of 79% of intervals ($SD = 10$). During superimposition, scratching substantially decreased during the second session, and became more variable during sessions two through nine. Scratching increased minimally during sessions 23 through 27 of this condition, but overall reached only an overall mean of 75% of intervals ($SD = 21$). The total number of reinforcers delivered was 2338 ($M = 93.5$; range, 29-150). During removal of superimposition, scratching decreased slightly during sessions 32 through 37 and then returned to baseline levels, with scratching occurring at an overall mean of 73% of intervals ($SD = 14$).

As seen in the bottom panel, Libby’s scratching was also variable throughout baseline, and occurred at a mean of 61% of intervals ($SD = 12$). Libby’s scratching increased minimally throughout superimposition, occurring at a mean of 74% of intervals ($SD = 13$). Variability remained similar to baseline during superimposition. The total number of reinforcers delivered per session was 2801 ($M = 93.4$; range, 56-128). During removal of superimposition, scratching initially decreased in a gradual and orderly progression. After 11 sessions, scratching became increasingly variable, exhibiting a pattern characteristic of an extinction burst. Following this, however, scratching returned to baseline levels ($M = 61%$, $SD = 19$). Low levels of problem behavior were briefly observed for both Rose and Libby in the removal of superimposition condition.
Figure 2. Treatment Evaluation Outcomes for Rose and Libby
Environmental enrichment. The results of the free-operant preference assessments for Rose and Libby are depicted in Figure 3. For Rose, Bumble Ball®, Chilla Chilla®, and keyboard were selected for the EE evaluation because they resulted in the lowest rates of scratching. The "cross-over" pattern across toys on the x-axis demonstrates the changing relation between scratching and toy engagement for Rose. As can be seen in Figure 2, EE was efficacious in decreasing scratching (M = 4.3%, SD = 2.8) and resulted in high and consistent levels of toy engagement (M = 98.3%, SD = 1.7). For Libby, musical finger painting, animal abacus, and impression board were selected for the EE evaluation because they resulted in the lowest rates of scratching. As can be seen in Figure 2, EE resulted in lower but variable levels of scratching (M = 23%, SD = 13) and variable levels of toy engagement (M = 49%, SD = 18.6).
Figure 3. Free-operant Preference Assessment Outcomes for Rose and Libby.
DISCUSSION

The current study assessed the functions of stereotypy of three children diagnosed with autism, and evaluated a tangible superimposition procedure as a reductive intervention with two of these children. Finally, EE was implemented as an alternative treatment when the superimposition procedure proved ineffective in reducing stereotypy with both children.

Results of the functional analyses are consistent with much research on stereotypy, in that the patterns of stereotypy for all participants were indicative of automatic reinforcement functions. Interestingly, each participant’s functional analysis resulted in each of the three different patterns typically interpreted as automatic reinforcement: 1) Doug’s hand flapping was differentially high in the no-interaction condition, 2) Libby’s scratching was high and variable across all conditions, and 3) Rose’s scratching was high in the no-interaction condition and low in the control condition (LeBlanc et al., 2000).

In the current study, a no-interaction condition was used rather than an alone condition because sessions were conducted at home and in school. A potential problem with using no-interaction or ignore conditions is that if the target behavior is maintained by social reinforcement, the presence of the experimenter may serve as an $S^D$ and evoke problem behavior during this condition. Although it is plausible that this could have occurred with Libby, programmed discriminative stimuli were matched to conditions to reduce the chance of this occurring (Conners et al., 2000).

The results of the superimposition procedure evaluation showed that any reductive effects of this procedure are limited and brief. These results were replicated
across participants who were homogenous on variables of topography of stereotypy, function of stereotypy, language skills, and adaptive skills. The decrease in variability during superimposition found in the Neisworth et al. (1985) investigation was not replicated in the current study, nor was this pattern observed in other studies on the superimposition procedure.

There were, however, some subtle differences between Rose and Libby’s responding. For example, during superimposition, the initial decrease in responding observed in the Wylie and Grossman (1988) study also occurred with Rose, but not with Libby. Wylie and Grossman, however, found that responding remained at this low level throughout superimposition, and speculated that this occurred due to food consumption. Rose, however, continued to engage in scratching while she consumed the food in later sessions of the superimposition phase. Anecdotally, delivering food appeared to disrupt scratching in that she began reaching for the food, touching the experimenter’s arm and saying “eh”. During superimposition, Rose’s scratching did not increase, and thus a reinforcement effect was not demonstrated. In other words, it cannot be determined from those data if tangibles came to control Rose’s scratching in any way. Although the occurrence of problem behavior during removal of superimposition may suggest extinction-induced aggression, it is also plausible that when tangibles were no longer delivered Rose began engaging in pinching responses, which may have been reinforced with tangible items in the past. The initial slight reduction in scratching during removal may have been due to time spent attempting to access food. Libby’s responding during superimposition and removal was slightly different than Rose’s. Minor increases in scratching did occur during superimposition, suggesting a possible reinforcement effect.
Further evidence for tangible control of scratching was observed during removal in the brief occurrence of problem behavior, the orderly decrease in trend, and increased variability during this phase. In addition, anecdotally, during removal Libby initially engaged in scratching and then immediately looked at the bag containing food. This was unusual behavior for Libby, who rarely displayed focused eye contact on any item. It is possible this effect was observed with Libby and not Rose because Libby received over 500 more pairings than Rose did (2801 and 2338, respectively).

The data obtained in the current study did not replicate other studies that observed robust effects of the superimposition procedure. There are several potential factors to consider in evaluating this discrepancy. First, there may have been differences in initial baseline rate of behavior that would differentially be affected by this procedure. However, Rose and Libby’s mean baseline level of stereotypy (79% and 61% of intervals, respectively) is comparable to that found in the Neisworth et al. (1985) investigation (70% and 52% of intervals). It is difficult, however, to make comparisons to other studies with positive results because data are reported in responses per session or per interval rather than percentage of intervals as in the current study. Second, it is possible that the superimposition procedure is more effective with target behaviors maintained by social reinforcement. This would account for the rapid changes in behavior observed across all participants in Schmid (1986), in that the off-task behaviors were likely previously maintained by social-negative reinforcement (escape). Third, it is possible that this procedure works differently with participants with different language and skill levels. Participants in all other studies were diagnosed with mental retardation and reportedly displayed limited language skills. Although specific information was not
reported, Rose and Libby may have functioned at a lower level than participants in previous studies. Fourth, the number of sessions of tangible delivery could certainly produce differential results. However, the 25 and 30 sessions used with Rose and Libby is comparable to other studies, and clearly more than used in Schmid in which the effect was observed. Fifth, as was previously mentioned, the robust findings in Schmid may have been due to the change in schedule in addition to type of stimulus delivered, essentially functioning as the procedure used in studies on the IRE. Sixth, the maintenance observed in Neisworth et al. may have been due to contingencies that maintained the decrease in the natural environment. For example, Matthew may have received reinforcement for engaging in behaviors other than hand flapping. Given the reported inverse relationship between stereotypy and behaviors such as appropriate play and increased acquisition of skills, it is possible that Matthew learned novel behaviors that were incompatible with hand flapping during that time. Last, it may be that the superimposition procedure does not produce a robust or reliable decrease in behavior. Given the data from evaluations of the superimposition procedure thus far, this may be the most plausible explanation for the inconsistent findings on this phenomenon. Although it may indeed occur, the conditions under which it occurs are not readily identified. Furthermore, given these results, the procedures used would likely not be considered practical or have high social validity. To be effective, tangibles would likely need to be delivered and subsequently withdrawn throughout the day, for several days, across settings. Given this extended period of “reinforcement” of stereotypy, treatment integrity by caregivers or teachers would presumably be low in the absence of a reliable and robust effect.
EE was efficacious in decreasing scratching for both participants. This finding is consistent with other studies utilizing hypothesis-based EE interventions (e.g., Piazza et al. 2000) and provides further support for its use in decreasing stereotypy for children diagnosed with autism. In the current study, the type of sensory stimulation maintaining scratching was not experimentally confirmed. However, the toys used did result in decreases in scratching. Following the study, the experimenter made recommendations to caregivers about the characteristics of toys that produce similar effects (e.g., toys that produce varied sounds and stimulation to the fingers). Rotation of toys was recommended to prevent satiation, especially with Libby, whose scratching remained variable when given access to these toys. Toy engagement was also variable, as Libby frequently scratched and engaged with toys simultaneously. For Rose, EE resulted in immediate and substantial decrease in scratching. Unlike Libby, Rose engaged with toys without scratching them or other objects at the same time. Interestingly, the pattern observed in the control condition of the functional analyses appeared to be predictive of the effectiveness of EE in reducing scratching for both Rose and Libby. When Rose was exposed to highly preferred toys and frequent noncontingent attention in the control sessions, she engaged in differentially lower levels of scratching ($M = 18.3\%$) than in all other conditions. Similarly, exposure to highly preferred toys during EE resulted in scratching at near zero levels ($M = 4.3\%$). The difference observed may be due to the fact that the toys present during EE were matched to hypothesized sensory consequences of stereotypy. Libby's responding during some control sessions was slightly lower ($M = 30.6\%$) than during other conditions, but overall was undifferentiated. Similarly, environmental enrichment resulted in low ($M = 23\%$) but variable rates of scratching.
Future research might further investigate the utility of functional analysis patterns in predicting the efficacy of EE.

In conclusion, the superimposition procedure resulted in only minor and temporary decreases in stereotypy for both participants. Although the investigation was arranged to employ more rigorous procedures than those used in previous studies, positive results from previous studies were not replicated. These data indicate that the reductive effects of the superimposition procedure are not a very reliable finding, nor are they particularly robust. If this procedure had reliably produced substantial decreases in stereotypy, it would have contributed to the current knowledge base on transfer of stimulus control and may have contributed to the development of a useful intervention for problem behavior. However, given that the effect was not robust or durable, and the procedures may not by themselves be practical for clinical intervention in the natural environment, utilization of this procedure is not currently recommended.
REFERENCES


stimuli during multielement functional analyses. Journal of Applied Behavior
Analysis, 33, 299-308.

examining the effects of extrinsic rewards on intrinsic motivation. Psychological
Bulletin, 125, 627-668.

of noncontingent access to single- versus multiple-stimulus sets on self-injurious

format for assessing reinforcer preferences. Journal of Applied Behavior
Analysis, 29, 519-533.

observers, mediators, and generalization programmers using an icing procedure.

Behavior Modification, 2, 371-385.


and side effects of overcorrection. Journal of Applied Behavior Analysis, 7, 385-
390.

Improvements in the behavior of children with autism following massage therapy.


Appendix

Approval Letter From the Human Subjects Institutional Review Board
Date: October 10, 2002

To: James Carr, Principal Investigator
    Tina Sidener, Student Investigator for thesis

From: Mary Lagerwey, Chair

Re: HSIRB Project Number 02-09-02

This letter will serve as confirmation that your research project entitled "Reducing Stereotypic Movements with Reinforcer Displacement" has been approved under the full category of review by the Human Subjects Institutional Review Board. The conditions and duration of this approval are specified in the Policies of Western Michigan University. You may now begin to implement the research as described in the application.

Please note that you may only conduct this research exactly in the form it was approved. You must seek specific board approval for any changes in this project. You must also seek reapproval if the project extends beyond the termination date noted below. In addition if there are any unanticipated adverse reactions or unanticipated events associated with the conduct of this research, you should immediately suspend the project and contact the Chair of the HSIRB for consultation.

The Board wishes you success in the pursuit of your research goals.

Approval Termination: September 18, 2003