Behavioral Therapy for Treatment of Stereotypic Movements in Nonautistic Children
Jonathan M. Miller, Harvey S. Singer, Dana D. Bridges and H. Richard Waranch

J Child Neurol 2006; 21; 119
DOI: 10.1177/08830738060210020701

The online version of this article can be found at:
http://jcn.sagepub.com/cgi/content/abstract/21/2/119

Published by:
SAGE Publications
http://www.sagepublications.com

Additional services and information for Journal of Child Neurology can be found at:
Email Alerts: http://jcn.sagepub.com/cgi/alerts
Subscriptions: http://jcn.sagepub.com/subscriptions
Reprints: http://www.sagepub.com/journalsReprints.nav
Permissions: http://www.sagepub.com/journalsPermissions.nav
Citations (this article cites 40 articles hosted on the SAGE Journals Online and HighWire Press platforms):
http://jcn.sagepub.com/cgi/content/refs/21/2/119
Motor stereotypies are patterned, coordinated, repetitive, involuntary movements that appear to be driven and nonfunctional, except as a form of self-stimulation. These movements can take many forms, including arm flapping, hand waving, finger wiggling, leg shaking, body rocking, and head bobbing. Although motor stereotypies have long been recognized to be common in children with autism, mental retardation, sensory deprivation, and neurodegenerative disorders, it has become increasingly evident that stereotypic movements are common in seemingly normal individuals.  

The decision to treat stereotypic movements is generally based on the severity and psychosocial impact of the disorder. Treatment for stereotypies has included both pharmacologic and behavioral approaches. A wide variety of drugs have been studied in children with autistic spectrum disorders. Benzodiazepines, α-adrenergic agonists, opiate antagonists, β-blockers, antiepileptic drugs, antipsychotic agents, selective serotonin reuptake inhibitors, and others have shown variable effects. Behavior therapy has also been attempted in autistic or retarded individuals to decrease the occurrence of stereotypic movements. Studies have suggested that the behavior therapy technique of habit reversal can be effective in both behavioral and organic disorders and is particularly useful for repetitive, stereotyped behaviors. Habit reversal typically consists of several components: the most essential are (1) awareness training, in which the patients learn voluntarily to exhibit the target behavior to help them become more aware of when it is occurring, and (2) competing response training, in which the patients learn to inhibit the target behavior by substituting another behavior that is incompatible with it.
inary evaluation of the efficacy of a behavioral protocol that combines modified habit reversal therapy with differential reinforcement of other behavior in a small sample of nonautistic children with physiologic stereotypies, including hand or arm flapping, waving, clenching, and head bobbing.

METHODS

Subjects
Ten boys and two girls, ranging in age from 6 to 14 years (mean = 10.2 years, SD = 2.27 years), participated in this study. All subjects were diagnosed with a physiologic stereotypic movement disorder based on the following criteria: (1) movements with a fixed pattern that are repetitive, purposeless, and rhythmic; (2) movements that are not habits, mannerisms, compulsions, or complex motor tics; (3) onset before age 2 years; (4) no reported premonitory urge; (5) temporary suppression of movements by an external stimulus or distraction; and (6) no evidence of autism, pervasive developmental disorder, or sensory deprivation. Ten children were evaluated in the Pediatric Neurology Clinic at Johns Hopkins Hospital for repetitive movements before being chosen for behavioral therapy. The other two subjects were referred directly to a behavioral psychologist (H.R.W.), and videotapes of their movements were reviewed by a pediatric neurologist (H.S.S.). All subjects were seen on a fee-for-service basis. None of the subjects were seeing a mental health practitioner or taking either psychotropic or tic-suppressing medication during treatment.

Behavioral Therapy

Session 1
Parents and children were seen together for a 60-minute initial consultation. A brief general history was obtained, baseline data were collected about each child’s stereotypic behaviors by using the measures described below, and a brief description of the behavioral treatment protocol to be implemented was provided to parents and subjects. The DuPaul’s attention-deficit hyperactivity disorder (ADHD) screening scale,22 the Child Global Assessment Scale,23 the Stereotypy Linear Analog Scale, and the Stereotypy Severity Scale were completed by parents. Habit reversal training was described and then practiced five times by asking subjects voluntarily to exhibit their stereotypic behavior. During these practice sessions, parents were instructed to take into consideration the dimensions of number, frequency, intensity, and interference in making their ratings. Parents were also instructed to estimate the amount of time during each 10-minute session that the stereotypy occurred and record this on a separate data sheet. Habit reversal training was then practiced as in session 1, and parents were instructed to continue with the habit reversal training at home.

Additional Sessions
Between sessions, parents were encouraged to call or e-mail Dr Waranch with questions and interim progress reports. The additional sessions took place at 2- to 4-week intervals and then were spaced at increasingly longer intervals. Compliance with both habit reversal training and differential reinforcement of other behavior practice was reviewed. For all of the subjects, adjustments were made with the differential reinforcement of other behavior training to include (1) gradually increasing the frequency of practice sessions to four per day, (2) increasing the length of the sessions from 10 to 20 minutes, and (3) including an additional, more tangible reinforcer when necessary (ie, a small amount of money, such as 25 cents, or points toward earning a desired prize, such as a computer game) for successful trials when no stereotypy was exhibited.

Auditory Feedback
For the two girls with head bobbing as the stereotypy, parents constructed barrettes with small paper tubes attached that contained either small bells or metal pellets. Subjects wore these barrettes during differential reinforcement of other behavior practice sessions and at other times to help them become more aware of their head bobbing.

Follow-Up Telephone Calls
The parents of all subjects were contacted for a progress report before completion of the study. Parents were asked in a structured interview format to comment on the current quality, frequency, duration, and situations that induce the stereotypic movement, as well as the presence of any new movements. Specific questions were asked about performance in school, the need for any special services (ie, special education), involvement in psychological or psychiatric therapy, any new medications or diagnoses for the child, and the level of motivation shown by the child for stopping the movement. Finally, the parents completed the Stereotypy Severity Scale, Child Global Assessment Scale, and Stereotypy Linear Analog Scale and were asked to give an estimate of percent improvement from baseline.

Measures

Stereotypy Severity Scale
The Stereotypy Severity Scale was modified from the Yale Global Tic Severity Scale and designed to measure stereotypic movements rather than tics.25 This consists of two components, one pertaining to the motor dimensions of the movement and the other pertaining to the global effects of the movement. The motor component rates the movement along four discriminate dimensions, including number (0–3 points), frequency (0–5 points), intensity (0–5 points), and interference (0–5 points), for a total of 18 points (Figure 1). An independent rating of global impairment caused by the movement (up to 50 points) is added to obtain the total score (maximum 68 points), which is elevated with worsening severity of stereotypic movements.

Stereotypy Linear Analog Scale
This consisted of a line 10 cm long on which the parents were asked to rank their child’s stereotypy during the past few days from 0 = the best it has ever been to 10 = the worst it has ever been by making a mark on the line. Parents were instructed to take into consideration the dimensions of number, frequency, intensity, and interference in making their ratings.
Figure 1: Components and point scoring breakdown of the Stereotypy Severity Scale.

Child Global Assessment Scale Scoring Guide

The Child Global Assessment Scale is a numeric assessment (maximum of 100 points) of the child's overall function at home, at school, and with peers with respect to activities and interests, school performance, coping ability, mood, anxiety, and antisocial acts. The higher the score, the better the overall function.²³

Percent Improvement from Baseline

Parents were asked to give a percent improvement in stereotypic movements from the child's baseline. A higher percentage correlates with a greater degree of improvement.

Motivation Ranking

Parents were asked to rank, from 1 to 4, their child's motivation to stop the movement at the onset of therapy (1 = not motivated; 2 = slightly motivated; 3 = moderately motivated; 4 = very motivated).

Statistical Analyses

Exploratory data analytic methods were employed to test for distributional assumptions of normality among the key outcome measures (Stereotypy Severity Scale, Stereotypy Linear Analog Scale, Child Global Assessment Scale, and percent improvement from baseline) at baseline and at telephone follow-up. These outcome measures were analyzed in two ways: (1) the means at the onset of therapy were compared with those at telephone follow-up to look for statistically significant differences, and (2) a single number summary of the change in score from baseline to telephone follow-up was constructed as the ratio of this change to the baseline score (known as the change ratio). This change ratio can be interpreted as the percent change in mean score from baseline.

In addition, the distributions of important stratification variables were tested, including the motivation ranking and number of treatments. Confirmatory data analytic methods were used to investigate the relationship between the number of treatments received and the outcome measures. Correlation between the change ratios and the number of treatments was assessed using a nonparametric correlation coefficient, Spearman's rho. Based on the distributional assessment of the number of treatments received, the participant cohort was dichotomized around three treatments. Thus, change in outcome measures was compared between those receiving fewer than three treatments and those receiving three or more treatments by using the Wilcoxon rank sum test. This test was also used to compare the change ratios of the outcome measures when the patients were divided into two groups based on level of motivation (very motivated versus less motivated). All statistical analyses were performed by using SAS, version 8.2 (SAS Institute, Cary, NC).

RESULTS

Demographic data on the 12 patients, including a description of their stereotypic movements, age at onset, age at the start of therapy, the number of treatment sessions, and the presence of any comorbidities, are presented in Table 1. The majority were male (83%). Ten had complex stereotypies (hand or arm flapping, wiggling, clenching, shaking), and two had head-bobbing movements. The mean age at onset of the stereotypic movements in this population...
was 16.6 months. The mean duration of behavioral therapy was 4.25 months, the mean age at the start of behavioral therapy was 9.2 years, and the mean age at the time of telephone follow-up was 10.2 years.

**Outcome Measures**

The main outcome measures, the Stereotypy Severity Scale, Stereotypy Linear Analog Scale, and Child Global Assessment Scale, with their values at the onset of therapy and at telephone follow-up are presented in Table 2. The Stereotypy Severity Scale and Stereotypy Linear Analog Scale showed significant improvement over baseline \( (P = .046 \text{ and } P = .009, \text{ respectively}) \), and the decline in mean score from baseline as described by the change ratio was 26% for the Stereotypy Severity Scale and 38% for the Stereotypy Linear Analog Scale. The Stereotypy Severity Scale was also divided into its two components, motor and global, with the motor component decreasing from a mean of 11.6 to a mean of 10.0, which showed a trend toward significance \( (P = .063) \), and the global component decreasing from 29.2 to 19.6, which was statistically significant \( (P = .012) \). The Child Global Assessment Scale showed a trend toward improvement \( (P = .286) \). The change ratio for the Child Global Assessment Scale exhibited a 10% improvement in mean score from baseline. The fourth outcome measure examined was the percent improvement from baseline. The mean percent improvement from baseline was 54.67%, with a standard error of 9.55.

**Number of Treatment Sessions Versus Outcome**

The number of treatment sessions was correlated with the outcome measures in two ways: (1) with treatment number as a continuous variable and (2) with treatment number stratified around three treatment sessions. When the correlation was examined as a continuous variable, the Stereotypy Severity Scale and Stereotypy Linear Analog Scale indicated clear trends toward a relationship with increasing number of treatments leading to a decline in score \( \text{Stereotypy Severity Scale } r = -.53, P = .074; \text{ Stereotypy Linear Analog Scale } r = -.64, P = .026 \). The Child Global Assessment Scale did not show a significant trend \( (r = .28, P = .385) \).

The patients were then divided into two groups based on the number of treatment sessions they attended. Five patients went to two sessions each, and the other seven patients went to three or more treatment sessions (range 3–10 sessions). Table 3 shows the change ratio for each outcome measure in both groups. The Stereotypy Severity Scale had a 6% decline in mean score from baseline for the two-session group and a 41% decline in the group attending more than two sessions, which is a trend toward significance \( (P = .144) \). The Child Global Assessment Scale exhibited a 1% increase in mean score from baseline in the two-session group and a 10% increase in the > two-session group. This difference shows a trend toward significance \( (P = .413) \). The Stereotypy Linear Analog Scale exhibited a 7% decrease in the two-session group and a 60% decrease in the > two-session group \( (P = .023) \). Finally, the percent improvement from baseline was 29% in the two-session group and 75% in the > two-session group, of borderline significance \( (P = .059) \).

**Motivation Versus Outcome**

The level of motivation of the patients at the onset of treatment was dichotomized. Three patients were rated at the highest level of motivation \( (4 = \text{very motivated}) \), and the other nine had less motivation. The change ratios for each outcome measure for both groups are shown in Table 4. For all of the outcome measures except the Child Global Assessment Scale, the motivated patients had a greater change. The Stereotypy Severity Scale showed a 58% decline in the motivated group and a 16% decline in the group with less motivation \( (P = .330) \). The Stereotypy Linear Analog Scale showed a 79% decline in the motivated group and a 24% decline in the less motivated group \( (P = .064) \), and the percent improvement from baseline was 88.7% in the motivated group and 43.8% in the less motivated group \( (P = .049) \). No significant difference was observed for the Child Global Assessment Scale between groups.

**DISCUSSION**

This study was undertaken to provide a preliminary assessment of the effectiveness of a behavioral treatment combining modified habit reversal therapy and differential reinforcement of other behavior in reducing the occurrence of stereotypic movements in a small group of nonautistic children. On the basis of known improvements with other motor abnormalities, such as tic disorders, skin picking, trichotillomania, stuttering, thumb sucking, self-destructive oral habits, pica, and nail biting, we hypothesized that behavioral therapy would be effective at decreasing the movements and their effect on the patients' overall function.

Analysis of the main outcome measures, the Stereotypy Severity Scale and Stereotypy Linear Analog Scale, showed significant improvement from baseline to telephone follow-up. On post hoc analysis of the Stereotypy Severity Scale, the motor and global components were both improved. Confirmation of the motor component of the Stereotypy Severity Scale was obtained by the corresponding improvement in the Stereotypy Linear Analog Scale because they are both measurements of the motor dimensions of the movement. The Child Global Assessment Scale, a measure of the children's overall function not specifically related to the movement, showed some trends toward improvement. This corroborated the improvement seen in the global component of the Stereotypy Severity Scale. Not surprisingly, the Stereotypy Severity Scale, Stereotypy Linear Analog Scale, and percent improvement from baseline showed a correlation between motivation or attendance at a higher number of treatment sessions and a reduction in stereotypic movements. These trends were best exemplified when the treatment number was stratified around three treatments, with patients receiving three or more treatments showing a trend toward better results.

This study was limited by the size of the population, variable age, gender, initial severity of the movement, comorbidities, and length of follow-up. Another limitation is the outcome measures used; except for the Child Global Assessment Scale, these scales have not been validated as measures of stereotypic movements and are, to some extent, subjective. Also, this study was neither controlled nor blinded. Despite these limitations, the results of this preliminary study are encouraging in demonstrating the effectiveness of behavior therapy in reducing stereotypic movements in nonautistic children.

Stereotypic movements have been hypothesized to have a “function” for autistic and mentally retarded persons, such as stress reduction, sensory reinforcement, arousal modulation, communication, protest, or escape. In contrast, biologic mechanisms involving dopamine and the striatum have also been hypothesized. The psychosocial effects of motor stereotypies...
### Table 1. Demographic Information on the 12 Patients in the Study, Including Their Stereotypic Movements, Age at Onset of the Movement, Age at Start of Therapy, Duration of Therapy, Comorbidities, Pregnancy or Delivery, and Development

<table>
<thead>
<tr>
<th>Patient</th>
<th>Sex</th>
<th>Stereotypy</th>
<th>Age at Onset (mo)</th>
<th>Age at Start of Therapy (yr)</th>
<th>Duration of Training Sessions (mo)</th>
<th>Number of Training Sessions</th>
<th>Time From Onset of Therapy to Telephone Follow-Up (mo)</th>
<th>Comorbidity</th>
<th>Pregnancy or Delivery</th>
<th>Development</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>M</td>
<td>Hand flapping</td>
<td>7</td>
<td>6</td>
<td>&lt; 1</td>
<td>2</td>
<td>11</td>
<td>None</td>
<td>Normal</td>
<td>Speech delay, motor delay (delayed walking), sensory integration disorder</td>
</tr>
<tr>
<td>2</td>
<td>M</td>
<td>Shirt/object squeezing</td>
<td>12</td>
<td>9</td>
<td>10</td>
<td>10</td>
<td>26</td>
<td>None</td>
<td>Normal</td>
<td>Normal</td>
</tr>
<tr>
<td>3</td>
<td>M</td>
<td>Hand flapping</td>
<td>12</td>
<td>6</td>
<td>6</td>
<td>4</td>
<td>6</td>
<td>ADHD</td>
<td>Normal</td>
<td>Normal</td>
</tr>
<tr>
<td>4</td>
<td>F</td>
<td>Head bobbing</td>
<td>12</td>
<td>11</td>
<td>8</td>
<td>6</td>
<td>14</td>
<td>Tics</td>
<td>Normal</td>
<td>Motor delayed (delayed standing and walking)</td>
</tr>
<tr>
<td>5</td>
<td>M</td>
<td>Finger wiggling, hand flapping</td>
<td>18</td>
<td>9</td>
<td>5</td>
<td>7</td>
<td>7</td>
<td>Tics</td>
<td>Normal</td>
<td>CP with right-sided hemiparesis, speech delay</td>
</tr>
<tr>
<td>6</td>
<td>M</td>
<td>Face rubbing, finger wiggling</td>
<td>30</td>
<td>10</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>Anxiety, nonspecific learning disability</td>
<td>Normal</td>
<td>Normal</td>
</tr>
<tr>
<td>7</td>
<td>M</td>
<td>Head shaking, hand movements, shoulder shrugging</td>
<td>24</td>
<td>9</td>
<td>1</td>
<td>2</td>
<td>6</td>
<td>Tics</td>
<td>Normal</td>
<td>Normal</td>
</tr>
<tr>
<td>8</td>
<td>M</td>
<td>Hand flapping with running</td>
<td>30</td>
<td>8</td>
<td>&lt; 1</td>
<td>3</td>
<td>15</td>
<td>None</td>
<td>Normal</td>
<td>Normal</td>
</tr>
<tr>
<td>9</td>
<td>M</td>
<td>Hand flapping, knee slapping</td>
<td>24</td>
<td>9</td>
<td>1</td>
<td>2</td>
<td>13</td>
<td>None</td>
<td>Normal</td>
<td>Normal</td>
</tr>
<tr>
<td>10</td>
<td>M</td>
<td>Hand clenching with facial grimace</td>
<td>12</td>
<td>9</td>
<td>11</td>
<td>8</td>
<td>13</td>
<td>Learning disability for reading comprehension</td>
<td>Normal</td>
<td>Normal</td>
</tr>
<tr>
<td>11</td>
<td>M</td>
<td>Arm and jaw extension with facial grimace</td>
<td>12</td>
<td>12</td>
<td>&lt; 1</td>
<td>2</td>
<td>14</td>
<td>Non-specific learning disability</td>
<td>Normal</td>
<td>Normal</td>
</tr>
<tr>
<td>12</td>
<td>F</td>
<td>Head bobbing</td>
<td>6</td>
<td>12</td>
<td>4</td>
<td>6</td>
<td>17</td>
<td>Depression</td>
<td>Normal</td>
<td>Normal</td>
</tr>
<tr>
<td>Mean</td>
<td>0.83</td>
<td>(male)</td>
<td>16.58</td>
<td>9.17</td>
<td>4.25</td>
<td>4.5</td>
<td>12.08</td>
<td>Depression</td>
<td>Normal</td>
<td>Normal</td>
</tr>
<tr>
<td>Standard error</td>
<td>0.11</td>
<td></td>
<td>2.43</td>
<td>0.56</td>
<td>1.15</td>
<td>0.81</td>
<td>5.91</td>
<td>Depression</td>
<td>Normal</td>
<td>Normal</td>
</tr>
</tbody>
</table>

ADHD = attention-deficit hyperactivity disorder; CP = cerebral palsy; C/S = cesarean section; NICU = neonatal intensive care unit.
are variable. In children requiring therapy, there are multiple pharmacologic options, but no formal studies have been performed in the nonautistic population. Habit reversal therapy combined with differential reinforcement of other behavior provides a potentially valuable treatment option. Implementation of this therapy requires a motivated patient, multiple training sessions, and an experienced therapist. Further study with a larger number of patients in a randomized, blinded protocol with longer follow-up is needed.

**Acknowledgments**

We thank Richard Skolasky for assistance with statistical analysis and Pamala Talalay for editorial comments.

**References**


**Table 2.** Comparison of Baseline Outcome Measures to the Measures at Telephone Follow-Up

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Initial Mean (SE)</th>
<th>Follow-Up Mean (SE)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stereotypy Severity Scale</td>
<td>40.46 (2.69)</td>
<td>29.58 (4.02)</td>
<td>.046</td>
</tr>
<tr>
<td>Child Global Assessment Scale</td>
<td>73.33 (3.76)</td>
<td>79.17 (4.17)</td>
<td>.286</td>
</tr>
<tr>
<td>Stereotypy Linear Analog Scale</td>
<td>5.95 (0.55)</td>
<td>3.25 (0.56)</td>
<td>.009</td>
</tr>
</tbody>
</table>

SE = standard error.

**Table 3.** Ratio* of the Change in Each Measure to the Baseline Value and Percent Improvement from Baseline for Groups Divided Based on Number of Treatment Sessions

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Two Sessions Mean (SE)</th>
<th>&gt; Two Sessions Mean (SE)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stereotypy Severity Scale</td>
<td>−0.06 (0.15)</td>
<td>−0.41 (0.10)</td>
<td>.144</td>
</tr>
<tr>
<td>Child Global Assessment Scale</td>
<td>0.01 (0.10)</td>
<td>0.16 (0.08)</td>
<td>.413</td>
</tr>
<tr>
<td>Stereotypy Linear Analog Scale</td>
<td>−0.07 (0.13)</td>
<td>−0.60 (0.13)</td>
<td>.028</td>
</tr>
<tr>
<td>Percent improvement from baseline</td>
<td>29.00</td>
<td>73.00 (7.45)</td>
<td>.059</td>
</tr>
</tbody>
</table>

SE = standard error.

*Change ratio = (value at telephone follow-up − value at baseline)/(value at baseline).

**Table 4.** Ratio* of the Change in Each Measure to the Baseline Value and Percent Improvement from Baseline for Groups Divided Based on Level of Motivation

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Less Motivated Mean (SE)</th>
<th>Motivated Mean (SE)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stereotypy Severity Scale</td>
<td>−0.16 (0.09)</td>
<td>−0.58 (0.20)</td>
<td>.139</td>
</tr>
<tr>
<td>Child Global Assessment Scale</td>
<td>0.09 (0.08)</td>
<td>0.12 (0.00)</td>
<td>.852</td>
</tr>
<tr>
<td>Stereotypy Linear Analog Scale</td>
<td>−0.24 (0.11)</td>
<td>−0.79 (0.21)</td>
<td>.064</td>
</tr>
<tr>
<td>Percent improvement from baseline</td>
<td>43.33 (9.79)</td>
<td>88.67 (9.35)</td>
<td>.049</td>
</tr>
</tbody>
</table>

SE = standard error.

*Change ratio = (value at telephone follow-up − value at baseline)/(value at baseline).


