Habit Reversal Treatment of Tic Disorders

A Methodological Critique of the Literature

JAMES E. CARR
IVY M. CHONG
Western Michigan University

One of the most commonly reported and successful behavioral interventions for tic disorders is habit reversal (HR). Several narrative literature reviews have adequately summarized the outcomes of these studies. The purpose of this article was to review studies that used HR to treat tics in terms of their methodological characteristics and rigor. Guidelines developed by the Task Force on Promotion and Dissemination of Psychological Procedures were used to evaluate the state of the literature. From an initial database that included 29 studies, 12 were included in the final analysis. Results indicate that although research has been conducted in this area for almost three decades, the majority of studies contain considerable methodological shortcomings. Based on the Task Force guidelines, the existing literature on the use of HR to treat tics can currently be classified as probably efficacious, and it almost meets the criteria for the well-established classification. Directions for future research are discussed.

Keywords: habit reversal; tic disorders; Tourette syndrome; empirically validated treatment

Tic disorders comprise a group of clinical syndromes that share primary characteristics (i.e., tic expression) and differ mainly in severity. The Diagnostic and Statistical Manual of Mental Disorders (4th ed.; DSM-IV; American Psychiatric Association, 1994) provides diagnostic criteria for three distinct tic disorders: Tourette’s disorder (a.k.a., Tourette syndrome; TS), chronic tic disorder, and transient tic disorder.1 TS is the most severe, and perhaps most common, tic disorder

AUTHOR’S NOTE: Address correspondence to James E. Carr, Department of Psychology, Western Michigan University, 1903 W. Michigan Ave., Kalamazoo, MI, 49008-5439; e-mail: jim.carr@wmich.edu.

BEHAVIOR MODIFICATION, Vol. 29 No. 6, November 2005 858-875
DOI: 10.1177/0145445505279238
© 2005 Sage Publications
and is characterized by multiple motor and phonic tics. Tics are typically defined as involuntary, sudden, and repetitive motor movements and phonic or audible expressions (Leckman & Cohen, 1988). Individuals diagnosed with tic disorders exhibit tics many times a day, across time in a fluctuating pattern. In addition, these individuals frequently meet diagnostic criteria for obsessive-compulsive disorder, attention-related disorders, and learning disabilities; however, tics are the only clinical behaviors that are identified in the DSM-IV diagnostic criteria for tic disorders.

The most common interventions for tic disorders are pharmacological (see Carpenter, Leckman, Scahill, & McDougle, 1999). Although several medications have been shown to produce varying degrees of success in reducing tics, behavioral treatments (e.g., contingency management, habit reversal [HR], negative practice, relaxation therapy, self-monitoring) have been frequently implemented in this area and are generally reported as producing favorable results (Peterson & Azrin, 1993). To date, there have been more than 100 published studies examining the efficacies of different behavioral treatments for individuals with tic disorders. In addition, four narrative literature reviews have been published in the area (Azrin & Peterson, 1988a; Houlihan, Hofschulte, & Patten, 1993; Peterson & Azrin, 1993; Turpin, 1983). One of the most prevalent and effective behavioral treatments identified in these reviews is HR (for a review, see Miltenberger, Fuqua, & Woods, 1998; Woods & Miltenberger, 1995).

In 1973, Azrin and Nunn developed HR to treat various habits (e.g., tics, nail biting, hair pulling). The original HR treatment package included nine components, including awareness training and competing-response training (these two are widely considered the most important components). In a typical HR session, participants are first instructed to vocally identify each tic, perhaps increasing their awareness of the tics, as in self-monitoring interventions (Wright & Miltenberger, 1987). Participants are then taught competing responses (Carr, 1995) to perform immediately after each tic for 1 min. to 3 min. In addition to these two components, relaxation therapy, self-monitoring, contingency management, and social support interventions are sometimes used as ancillary treatment components. However, several studies (e.g., Miltenberger, Fuqua, & McKinley,
1985; Woods, Miltenberger, & Lumley, 1996) have shown that simplified forms of HR that include only awareness training and competing-response training can be effective in reducing tics as well as other habit behaviors.

The literature documenting the effects of HR in the treatment of tics includes several dozen studies dating back to the original Azrin and Nunn (1973) investigation. The majority of these studies employed within-subject (i.e., single-case) design methodology to evaluate treatment effects with relatively few participants. Although the scientific logic underlying this approach is sound, low ns have been cited as a shortcoming of this literature (e.g., King, Scahill, Findley, & Cohen, 1999; Piacentini & Chang, 2001), often by researchers who typically use between-subjects designs with many participants. The original purpose of this article was to conduct a quantitative meta-analysis of this literature using techniques specifically designed for single-case design data (see Faith, Allison, & Gorman, 1996) in an effort to combine the effects of many studies. Thus, each participant’s data would be converted into a single data point to represent effect size in an aggregate analysis of the efficacy of HR. However, during our initial review of the literature, it became apparent that the quality of the studies in the area was quite variable. There were considerable discrepancies in the numbers of participants, objectivity of data collection, rigor of research designs, etc. Consequently, we felt it unwise to quantify such a methodologically diverse literature. The focus of our efforts then shifted to documenting this diversity in a more traditional narrative review that focused primarily on methodology instead of outcome. Although we would now be able to document the specific methodological differences between studies, we had no framework with which to evaluate the overall state of the literature.

In the 1993, Division 12 (Clinical Psychology) of the American Psychological Association commissioned a task force to develop guidelines with which psychological treatments could be evaluated for their level of empirical support in the scientific literature. This effort was primarily undertaken to help guide the dissemination of evidence-based interventions. In 1995, the Task Force on Promotion and Dissemination of Psychological Procedures published consensus-developed criteria for empirically validated treatments.
Within the Task Force’s framework, a treatment can be classified as either *well established* or *probably efficacious* based on the number and types of published studies that exist to support it.

Two different criteria can result in a classification of *well-established* treatment. The first requires at least two between-subjects–design studies demonstrating (a) a superior treatment effect compared to another treatment or a placebo condition or (b) equivalence to an already established treatment. The second requires at least nine single-case–design studies that have demonstrated a superior treatment effect compared to another treatment or a control condition. Studies meeting the above criteria should be produced by at least two independent research groups. In addition, treatments should be implemented with manuals and the characteristics of participants and samples should be clearly described. *Probably efficacious* treatments should have (a) at least two studies showing superior effects compared to a wait-list control group, (b) one or more studies meeting the *well-established* treatment criteria (but from only one research group), or (c) at least three single-case–design studies that meet the *well-established* treatment criteria. HR is currently classified as a *probably efficacious* treatment by the Task Force on Promotion and Dissemination of Psychological Procedures for the treatment of nail biting (Azrin, Nunn, & Frantz, 1980a) and thumb sucking (Azrin, Nunn, & Frantz-Renshaw, 1980); however, HR has not yet been classified for the treatment of tics.

Despite the worthy goal attempted by the Task Force’s classification system and the considerable effort that was evident in its development, the system does have its problems. For example, the system strongly favors the randomized clinical trial, which itself has been the center of some controversy (see Seligman, 1995). However, despite its flaws and controversy (see Elliot, 1998), the classification system has rapidly been accepted by the scientific community. For example, the system has recently been used to evaluate empirical research in such diverse areas as pediatric feeding problems (Kerwin, 1999), attention-deficit hyperactivity disorder (Pelham, Wheeler, & Chronis, 1998), clinical hypnosis (Lynn, Kirsch, Barabasz, Cardena, & Patterson, 2000), and chronic pain (Wilson & Gil, 1996). We adopted the Task Force guidelines to provide the final evaluation of our narrative review.
for two reasons. First, it is the only systematic classification system for clinical interventions that is based on existing empirical evidence. Second, one of the Task Force’s primary missions was to assist with intervention dissemination, which has been a criticism of the HR literature (Miltenberger, Fuqua, & Woods, 1998). Because the mainstream psychological and psychiatric communities address the evaluation and treatment of tics apart from other habits, the focus of our review is specifically on the treatment of tics.

METHOD

LITERATURE SEARCH

The goal of our literature search was to identify all of the experimental studies that used HR (or a variation thereof) to treat tics. Articles were selected for our review based on computerized searches of the PsycINFO and Medline databases using the following search criteria: HR and tics or Tourette, competing response and tics or Tourette. In addition, the reference lists of the articles identified through PsycINFO and Medline were reviewed for additional articles. Among these were two studies that evaluated simplified forms of HR but referred to them as self-administered overcorrection (Gross & Mendelson, 1982; Ollendick, 1981). The initial list included 29 articles that reported the use of HR to treat tics. Of these initial 29 articles, 5 non-English-language studies (Aguilar & Boehrt, 1983; Araki & Nakai, 1990; Araki & Okuma, 1985; Bados Lopez & Olle, 1984; Cloutier, 1985) were excluded from our review because of difficulties in accessing and translating the articles. In addition, we excluded three case studies (O’Brien & Brennan, 1979; Young & Montano, 1988; Zikis, 1983) and one report that did not provide sufficient information regarding the independent variable or experimental design (O’Connor, Gareau, & Blowers, 1993). In total, 20 studies remained in the database for methodological analysis according to the areas described below.
CATEGORIES

Each of the 20 studies that were included in the database (see Table 1) was evaluated for the information in the following categories.

Participant characteristics. The number of participants who exhibited tics was recorded for each study. In some studies (e.g., Azrin & Nunn, 1973), participants engaged in a variety of behaviors including tics; however, only participants who exhibited tics were included in our analysis. In addition, the age range and gender of these participants were recorded. Finally, we recorded whether participants exhibited motor tics, phonic tics, or both.

Dependent measures. Most studies in the database employed within-subject designs with repeated measures to evaluate treatment effects. Therefore, we recorded the unit of time that comprised the data point in each study. In other words, we recorded the length of time the participant was observed for data collection purposes (i.e., each session or observation period). We also recorded whether the primary dependent measure was direct and objective. A direct-observation measure was defined as any systematic data collection procedure in which an observer (other than the participant) recorded tics as they occurred (Carr & Rapp, 2001). Common examples of direct measures were event (i.e., a frequency count) and partial-interval recording. Finally, we recorded whether studies reported interobserver agreement (IOA) data for their dependent measures.

Experimental design. We recorded whether the experimental design employed in each study was a between-subjects or within-subject design. If a study employed a between-subjects design, its specific form (e.g., pretest-posttest control) was identified. Similarly, if a study used a within-subject design, its specific form (e.g., multiple baseline) was identified.

Follow-up assessment. We recorded whether each study included an assessment of treatment maintenance after the primary treatment phase had ended.
Treatment integrity. We determined whether each study included a measure of treatment integrity. A treatment integrity measure was defined as any systematic effort to determine whether the treatment was being implemented as designed, including the use of treatment manuals. Integrity is especially important to the HR treatment of tics because of the participant-implemented competing responses. Until recently, most of the studies in this area that assessed treatment integrity focused on the behaviors of the therapists (i.e., whether the therapists followed a training protocol) instead of those of the participants (i.e., treatment adherence or compliance). Because of the general paucity of these measures in this literature, we recorded any attempt regardless of its focus (i.e., therapists or participants).

Reductions from pretreatment assessment. We recorded whether the authors reported a positive treatment gain for the participants. A treatment gain was defined as any reduction in tic frequency from baseline or pretreatment assessment periods. A more objective measure (e.g., the percentage of nonoverlapping data statistic; Scruggs & Mastropieri, 1998) was not used because of the methodological variability of the studies (e.g., data-point length, objective versus subjective measures). We refer readers to Peterson and Azrin (1993) for quantitative summaries of the reductions reported in many of the studies listed in our database.

Social validity. We recorded whether each study included an assessment of social validity, which was defined as any formal attempt to assess treatment acceptability or outcome with direct (e.g., participants) or indirect consumers (e.g., parents). Examples of social validity assessments included participants completing posttreatment questionnaires (e.g., Miltenberger, Fuqua, & McKinley, 1985) and university students providing social-perception ratings of videotaped samples (e.g., Woods et al., 1996).

RESULTS

In the 20 studies that comprised our database, 114 participants were exposed to some variation of HR. The component common to all of
these HR variations was the performance of the competing response. Of these 114 participants, 108 were successfully treated, as reported by the authors. On the surface, a 95% efficacy rate with more than 100 participants appears to strongly support the use of HR for tic reduction. However, of these 20 studies, only 15 employed an experimental design capable of demonstrating experimental control of the independent variable. The 5 studies that employed A-B designs were excluded from our final analysis because the designs do not "permit a full experimental analysis of the controlling effects of the treatment[s] inasmuch as [their] correlative properties are quite apparent" (Barlow & Hersen, 1984, p. 142). In addition, the characteristic waxing and waning of many tics over time makes them particularly problematic for A-B designs. Of the 15 studies that employed experimental designs, only 12 of them incorporated direct measurement of the dependent variables. The 3 studies that did not include direct tic measurement were excluded because of the validity problems associated with self-monitoring in the absence of corroborative measures (Poling, Methot, & LeSage, 1995).

The analysis that follows is based on the remaining 12 studies that incorporated both experimental designs and objective data collection procedures. These studies included 90 participants (age range: 6 to 66), approximately 73% of which were treated in three studies (Azrin & Peterson, 1990; Miltenberger, Fuqua, & McKinley, 1985; O’Connor, Brault, et al., 2001). Approximately 66% of the participants were males and 34% were females. Of the participants, 71 (79%) exhibited motor tics, four (4%) of the participants exhibited phonic tics, and 15 (17%) of the participants exhibited both motor and phonic tics.

All 12 of the studies reported IOA data. The most common data-point (i.e., observation period) length was measured in minutes. Azrin and Peterson (1989) and Peterson and Azrin (1992) reported the shortest duration data points at 1 min and 2.5 min, respectively. The longest data point (i.e., 1 month) was reported by Azrin and Peterson (1990). The most common experimental design was the multiple-baseline design, which was implemented across behaviors, participants, and settings in 8 studies. The second most common experimental designs were the reversal \((n = 3)\) and alternating treatments \((n = 3)\).

Of the 90 participants included in the 12 studies, reductions from pretreatment assessment were reported for 85 (94%) of them. Impressively, 10 of the 12 studies included some follow-up or long-term assessment (mean = 10 months; range, 1 month to 2 years), most of which were successful. Only 7 of the studies reported any measure of treatment integrity, and only 3 of these focused on the behaviors of the participants (i.e., treatment adherence). Seven of the studies reported assessments of social validity, most of which were positive.

DISCUSSION

In summary, out of a database containing 20 studies, only 12 included experimental designs and objective measurement of dependent variables. Of these 12 studies, 94% of the participants experienced tic reductions, most of which maintained at long-term follow-up assessments. However, the studies were inconsistent in their evaluation of social validity and treatment integrity. The most methodologically sound study in our final database was reported by Woods et al. (1996). In this study, the authors exposed four children to simplified variations of HR, which resulted in successful tic reduction. From a methodological perspective, the authors used a multiple-baseline design to demonstrate experimental control, included direct tic measurement, and rigorously assessed IOA, treatment integrity, social validity, and long-term maintenance of treatment gains.

To be considered a well-established treatment according to the aforementioned Task Force guidelines, a successful treatment effect must be demonstrated with at least two between-subjects designs or nine within-subject designs. Of the 12 studies that were analyzed in the current article, 2 used a between-subjects design (Azrin & Peterson, 1990; O’Connor, Brault, et al., 2001). Of the 10 within-subject design experiments in the database, 9 reported successful results. Therefore, both of the design options that can lead to a well-established classification were met. In addition, the Task Force guidelines require studies to be produced by at least two research groups,
which was confirmed by our analysis. Finally, the guidelines require the use of manuals in treatment-outcome research. Although only one of the studies in our database specifically reported the use of a manual (O’Connor, Brault, et al., 2001), with several others referring to treatment protocols, we believe the rigorous assessment of treatment integrity that focuses on both therapist and participant behavior to be a reasonable substitute. Unfortunately, the vast majority of the studies in our database did not rigorously assess the implementation of their independent variables. In addition, this literature has been criticized for insufficient description of participant characteristics (e.g., King et al., 1999), which is also a Task Force requirement. Detailed participant descriptions are particularly important in this area given the comorbidity of tics with other psychological disorders (e.g., obsessive-compulsive disorder). Therefore, in our opinion, the current literature on the use of HR to treat tics technically does not support a classification of well established. Although the minimum numbers of successful independent studies exist, insufficient use of treatment manuals and assessment of treatment integrity, combined with insufficient information on participant characteristics, precludes the classification.

Although the literature might not currently support the Task Force’s most prominent classification, it does meet the requirements for a probably efficacious classification. However, even this classification might be considered relatively liberal for at least two reasons. First, we included studies that evaluated any variation of HR, not just the original treatment package; thus, the body of interventions was quite heterogeneous. These variations ranged from simplified HR (i.e., awareness training, competing-response training; Woods et al., 1996) to HR combined with cognitive therapy (O’Connor, Brault, et al., 2001). However, the effectiveness of these variations could also be interpreted as evidence of the generality of HR in the treatment of tics. Second, because most (10/12) of the studies we evaluated were successful reports using within-subject designs, there is an implicit publication bias (i.e., low-n failures may not have been published).

The probably efficacious classification is quite surprising after almost three decades of research on the use of HR for tic reduction. Although our initial database of 29 studies was reduced to 12 for the
final analysis, our intention is not to diminish the studies that were not included in our final database. Indeed, the original article by Azrin and Nunn (1973) was excluded for methodological reasons. However, we believe the role these studies best serve is to provide less formal convergent evidence (albeit not via rigorous experimental demonstration) for the use of HR to treat tics.

Recently, Miltenberger, Fuqua, and Woods (1998) called for wider dissemination of HR procedures. We enthusiastically agree with the authors. However, given the results of our methodological analysis, we believe that there is still a need for more rigorous demonstrations of HR for the treatment of tics. At least according to Task Force recommendations, there is a clear need for more studies that include direct measurement, experimental designs, detailed information about participants (including existing medications, which are common in individuals with tics), and evaluation of treatment adherence along with the use of treatment manuals (e.g., Woods, 2001). There is also a need for more frequent and direct evaluation of social validity to document the qualitative outcomes of HR techniques. Most studies assess only the frequency of tics, often to the exclusion of their severity, which could be an equally important behavioral dimension. In the absence of direct information on tic severity, the assessment of social validity would be even more valuable because it might reveal negative changes in severity. These recommendations are not offered solely for HR to achieve a more prominent Task Force classification but also to improve the overall quality of research in this area. Fortunately, a review of the studies listed in Table 1 indicates a trend toward methodological refinement among the more recent studies (e.g., Clarke, Bray, Kehle, & Truscott, 2001; O’Connor, Brault, et al., 2001; Woods et al., 1996).

The majority of studies we reviewed evaluated HR to treat motor tics, with few studies examining the effects on phonic tics. Given the topographical differences between competing responses for motor (e.g., head shaking) versus phonic (e.g., coprolalia) tics, we believe there is a clear need for more research on phonic tics. Furthermore, although several studies evaluated the effects of HR on both motor and phonic tics, these studies generally assessed the effects in a combined fashion, possibly obscuring differential responsiveness. Therefore, in
<table>
<thead>
<tr>
<th>Authors</th>
<th>Year</th>
<th>n</th>
<th>Age Range</th>
<th>Gender</th>
<th>Motor or Phonic Tics</th>
<th>Data-Point Length</th>
<th>Direct-Observation Measurement</th>
<th>Interobserver Agreement</th>
<th>Experimental Design</th>
<th>Follow-Up</th>
<th>Treatment Integrity</th>
<th>Reductions from Pretreatment Assessment</th>
<th>Social Validity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Azrin and Nunn</td>
<td>1973</td>
<td>4*</td>
<td>14-64</td>
<td>2 M, 2 F</td>
<td>Motor</td>
<td>1 day</td>
<td>No</td>
<td>No</td>
<td>A-B</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Azrin, Nunn, and Frantz</td>
<td>1980b</td>
<td>10</td>
<td>10-62</td>
<td>8 M, 2 F</td>
<td>Motor</td>
<td>1 day</td>
<td>No</td>
<td>No</td>
<td>Between subjects (Standard treatment control with repeated measures)</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Azrin and Peterson</td>
<td>1988b</td>
<td>3</td>
<td>28-42</td>
<td>3 M</td>
<td>Both</td>
<td>1 day, 1 week, 1 month</td>
<td>Yes</td>
<td>Yes</td>
<td>A-B</td>
<td>No**</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Azrin and Peterson****</td>
<td>1989</td>
<td>1</td>
<td>9</td>
<td>1 F</td>
<td>Motor</td>
<td>1 min, 1 day, 1 week</td>
<td>Yes</td>
<td>Yes</td>
<td>Reversal</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Azrin and Peterson****</td>
<td>1990</td>
<td>10</td>
<td>6-36</td>
<td>7 M, 3 F</td>
<td>Both</td>
<td>1 month</td>
<td>Yes</td>
<td>Yes</td>
<td>Between subjects (Waiting list control with repeated measures) + Multiple baseline (across participants)</td>
<td>No**</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Carr and Bailey****</td>
<td>1996</td>
<td>1</td>
<td>9</td>
<td>1 M</td>
<td>Phonics</td>
<td>5 min</td>
<td>Yes</td>
<td>Yes</td>
<td>Alternating treatments + reversal</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Carr, Bailey, Carr, and Coggin</td>
<td>1996</td>
<td>2</td>
<td>12</td>
<td>2 M</td>
<td>Both</td>
<td>5 min</td>
<td>Yes</td>
<td>Yes</td>
<td>Alternating treatments (treatments)</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Clarke, Rep, Kahle, and Truscott</td>
<td>2001</td>
<td>4</td>
<td>11-16</td>
<td>4 M</td>
<td>Both</td>
<td>18 min</td>
<td>Yes</td>
<td>Yes</td>
<td>Multiple baseline (across participants)</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes (3/4)</td>
<td>Yes</td>
</tr>
<tr>
<td>Finney, Rapoff, Hall, and Christophersen****</td>
<td>1983</td>
<td>2</td>
<td>11-12</td>
<td>2 M</td>
<td>Motor</td>
<td>10 min</td>
<td>Yes</td>
<td>Yes</td>
<td>Multiple baseline (Across behaviors and participants)</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

(continued)
### TABLE 1 (continued)

<table>
<thead>
<tr>
<th>Authors</th>
<th>Year</th>
<th>n</th>
<th>Age Range</th>
<th>Gender</th>
<th>Motor or Phonic Tics</th>
<th>Data Point Length</th>
<th>Direct Observation Measurement</th>
<th>Interobserver Agreement</th>
<th>Experimental Design</th>
<th>Follow-Up</th>
<th>Treatment Integrity</th>
<th>Pretreatment Assessment</th>
<th>Reductions from Pretreatment Assessment</th>
<th>Social Validity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fuata and Griffiths</td>
<td>1992</td>
<td>1</td>
<td>28</td>
<td>M</td>
<td>Phonic</td>
<td>10 min</td>
<td>Yes</td>
<td>No</td>
<td>A-B</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Gross and Mendelson</td>
<td>1982</td>
<td>1</td>
<td>11</td>
<td>M</td>
<td>Motor</td>
<td>1 day</td>
<td>Yes</td>
<td>Yes</td>
<td>A-B</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Miltenberger and Fuqua</td>
<td>1985</td>
<td>1*</td>
<td>67</td>
<td>F</td>
<td>Motor</td>
<td>1 day</td>
<td>No</td>
<td>No</td>
<td>Multiple baseline</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Miltenberger, Fuqua, and McKinley***</td>
<td>1985</td>
<td>9</td>
<td>9-60</td>
<td>M, F</td>
<td>Motor</td>
<td>1-30 min</td>
<td>Yes</td>
<td>Yes</td>
<td>Multiple baseline</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>O'Connor, Betz, et al.</td>
<td>2001</td>
<td>47*</td>
<td>18-60***</td>
<td>M, F</td>
<td>Both</td>
<td>10 min</td>
<td>Yes</td>
<td>Yes</td>
<td>Between subjects</td>
<td>No</td>
<td>Yes</td>
<td>Yes (8/9)</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>O'Connor, Gareau, and Borgeat</td>
<td>1997</td>
<td>7</td>
<td>23-49</td>
<td>M, F</td>
<td>Motor</td>
<td>unknown</td>
<td>No</td>
<td>No</td>
<td>Between subjects</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Peterson and Azrin***</td>
<td>1992</td>
<td>6</td>
<td>10-40</td>
<td>M</td>
<td>Both</td>
<td>2.5 min</td>
<td>Yes</td>
<td>Yes</td>
<td>Alternating treatments</td>
<td>No</td>
<td>No</td>
<td>Yes (5/6)</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Pray, Kramer, and Lindskog</td>
<td>1986</td>
<td>1</td>
<td>9</td>
<td>M</td>
<td>Motor</td>
<td>1 day</td>
<td>Yes</td>
<td>Yes</td>
<td>A-B</td>
<td>No</td>
<td>Yes</td>
<td>Yes (5/6)</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Sharenow, Fuqua, and Miltenberger***</td>
<td>1989</td>
<td>2</td>
<td>32-66</td>
<td>M, F</td>
<td>Motor</td>
<td>10 min</td>
<td>Yes</td>
<td>Yes</td>
<td>Multiple baseline</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Woods, Miltenberger, and Lumley***</td>
<td>1996</td>
<td>4</td>
<td>8-12</td>
<td>M, F</td>
<td>Motor</td>
<td>20 min</td>
<td>Yes</td>
<td>Yes</td>
<td>Multiple baseline</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

* Only participants who received habit reversal for tics were included.
** Follow-up assessments were implicit in the extended time frame in which these studies were conducted.
*** These data include participants who did not receive treatment. Demographic information on only the 47 participants who received treatment was not available in the article.
**** These 12 studies were included in our final methodological analysis.
addition to further research on the treatment of phonic tics, we recommend researchers evaluate the effects of HR across motor and phonic tics separately.

Given the contemporary behavioral literature’s focus on function-based interventions, it is striking that none of the studies in our database reported the use of pretreatment functional assessment. There is limited evidence to suggest that tics can serve operant functions (e.g., Carr, Taylor, Wallander, & Reiss, 1996; Malatesta, 1990; Watson & Sterling, 1998). However, HR appears to be effective irrespective of knowledge of tic function. Future research might address this by determining HR effects with tics of varying functions or by conceptualizing HR as a default intervention when social-reinforcement functions are disconfirmed during functional assessment.

NOTES

1. Because there is no evidence to suggest that transient tic disorder, chronic tic disorder, and TS have different etiologies or that they differentially respond to treatment, we reviewed studies whose primary purpose was to reduce clinically problematic tics. Thus, we do not differentiate between these studies on the basis of their participants’ diagnoses.

2. For more thorough discussions of treatment integrity and HR, we refer the reader to Carr, Bailey, Carr, and Coggin (1996) and Miltenberger, Fuqua, and Woods (1998).

3. It should be noted that the summary statistics in the Results section include data from the O’Connor et al. (2001) study, which reported group instead of individual data. Therefore, these summary statistics, especially those regarding treatment efficacy, should be viewed in a somewhat liberal manner.

REFERENCES


James E. Carr, Ph.D., is an associate professor of psychology and chair of the behavior analysis program at Western Michigan University. His current clinical and research interests include functional assessment and treatment of aberrant behavior, verbal behavior, noncontingent reinforcement, reinforcer identification, and assessment and treatment of tic disorders.
Ivy M. Chong, Ph.D., is currently the clinical manager of CARE Center for Autistic Children at Beaumont Hospitals. Her current clinical and research interests include the behavioral treatment of autism, verbal behavior, and functional assessment and treatment of problem behavior.