Accepted Manuscript

The efficacy of habit reversal therapy for tics, habit disorders, and stuttering: A meta-analytic review

Karina S. Bate, John M. Malouff, Einar T. Thorsteinsson, Navjot Bhullar

PII: S0272-7358(11)00075-4
DOI: doi: 10.1016/j.cpr.2011.03.013
Reference: CPR 1159

To appear in: *Clinical Psychology Review*

Received date: 19 September 2010
Accepted date: 29 March 2011

Please cite this article as: Bate, K.S., Malouff, J.M., Thorsteinsson, E.T. & Bhullar, N., The efficacy of habit reversal therapy for tics, habit disorders, and stuttering: A meta-analytic review, *Clinical Psychology Review* (2011), doi: 10.1016/j.cpr.2011.03.013

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

The Efficacy of Habit Reversal Therapy for Tics, Habit Disorders, and Stuttering: A Meta-Analytic Review

Karina S. Bate[1], John M. Malouff[1], Einar T. Thorsteinsson[1], & Navjot Bhullar

University of New England, Australia

Karina S. Bate, John M. Malouff, Einar T. Thorsteinsson, & Navjot Bhullar, Psychology, University of New England, Armidale, NSW, Australia.

[1] Karina S. Bate, John M. Malouff, and Einar T. Thorsteinsson contributed equally to this publication.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Abstract

A meta-analysis based on 575 participants in 18 studies found Habit Reversal Therapy (HRT) to be an efficacious intervention for a wide variety of maladaptive repetitive behaviors, including stuttering, tics, nail biting, temporomandibular disorder, thumb sucking, and mixed repetitive oral-digital habits. Compared to control conditions, HRT showed a large effect size pre-treatment to final post-treatment assessment, $d = 0.80$. Moderator analyses revealed significant treatment effects for HRT for most moderator levels, indicating that HRT is efficacious in a number of variations for a variety of types of maladaptive behaviors, across a wide range of sample characteristics. The findings provide substantial support for the efficacy of HRT for disorders it is commonly used to treat. The findings are consistent with recent arguments for the classification of HRT as a well-established treatment for tic and habit disorders.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

The Efficacy of Habit Reversal Therapy for Tics, Habit Disorders, and Stuttering: A Meta-Analysis

Habit reversal therapy (HRT) is commonly used to treat tic disorders, including Tourette's Syndrome (Woods & Miltenberger, 1995); habit disorders that involve trichotillomania, nail biting, temporomandibular disorder, thumb sucking, and repetitive oral-digital habits such as finger and lip biting (Keuthen, Bohne, Himle, & Woods, 2005); and stuttering (Conelea, Rice, & Woods, 2006). These disorders all involve repetitive, body-focused, unwanted behavior that causes distress or significant functional impairment to an individual (see Himle, Flessner, Bonow, & Woods, 2006). Miltenberger (2008) characterized the group of problems as habit behaviors, each with clear potential for development of a competing muscle movement.

HRT is a multicomponent behavioral intervention designed to reduce the manifestations of habit-based disorders. Conceptualized by Azrin and Nunn (1973), the original HRT package included 11 major techniques organized in five phases: (a) awareness training, (b) relaxation training, (c) competing response training, (d) motivation procedures, and (e) generalization training.

Awareness training is based on the premise that increased awareness of behavior facilitates self-control (Piacentini & Chang, 2005). A client is taught to discriminate each occurrence of the habit, including the environmental antecedents that precede its manifestation. Awareness training components include: (a) response description: identification of occurrences surrounding performance of the target behavior, (b) response detection: self-observation and diary monitoring of the topography, frequency and severity of target movements, (c) early warning: focusing on the sensations preceding behavior onset, and (d) situation awareness: acknowledgement of high-risk situations for the target behavior.

ACCEPTED MANUSCRIPT

Relaxation training is based on the premise that increases in stress lead to increases in habit behaviors (Piacentini & Chang, 2005). However, some experts have noted that the effects of relaxation training on the occurrence of target behavior have not been observed to generalize outside treatment sessions (Adams, Adams & Miltenberger, 2008; Piacentini & Chang, 2005).

Competing response training has the aim of rendering performance of the undesirable behavior impossible due to the use of antagonist muscles to the behavior (O'Connor, 2005a). This antagonistic response serves to normalize the behavior by replacing it with a less conspicuous action. Azrin and Nunn (1973) proposed that the competing response should (a) be opposite to the target behavior, (b) be able to be maintained for several minutes (Woods & Miltenberger, 1995), (c) produce heightened awareness of the behavior though muscle contraction, (d) be socially inconspicuous, (e) be compatible with normal activity, and (f) should strengthen the muscles antagonistic to the target behavior (Piacentini & Chang, 2005). Competing responses include the following: For tics, isometric tensing of muscles opposite to the tic movement (Azrin & Peterson, 1990); for stuttering, breathing slowly, deeply, and regularly, relaxing chest and throat muscles, and speaking upon exhaling (Waterloo & Gotestam, 1988); for thumbsucking, clenching both fists with the thumb inside (Christianson & Sander, (1987). Adams et al. (2008) and Miltenberger (2008) describe other competing responses for various repetitive behavior disorders.

Motivation training techniques are used to enhance the acceptability of HRT to children and their families (Piacentini & Chang, 2005) and to encourage treatment compliance. The three standard motivational techniques are (a) a habit inconvenience review that involves identification of the impairments, problems and discomforts caused by the habit, (b) social support training involving the recruitment of a person in the client's life to provide encouragement and support (e.g., prompts to use the competing response) in order to aid

ACCEPTED MANUSCRIPT

implementation of the procedures (Woods, 2006), and (c) public display, involving public performance and demonstration of the client's control of the target behavior.

Generalization training involves symbolic rehearsal in which the client imagines successful control in situations identified during the awareness stage (Woods & Miltenberger, 1995; Adams et al., 2008; O'Connor, 2005a). With each situation the client imagines beginning the target behavior, stopping the behavior, and performing the competing response (Miltenberger, Fuqua, & Woods, 1998).

Azrin and Nunn (1973) tested the original treatment procedure for HRT in just one treatment study. Contemporary applications of HRT follow this precedent and are typically implemented in a small number of sessions, followed by booster sessions as required (Adams et al., 2008). Woods (2006) described a short form of HRT as consisting of three sessions. The first session is devoted to clinical assessment of the client's intellectual, psychological and social functioning, and habit severity, typically in a 2 to 3 hour period. The primary objective of session two is the implementation of HRT. Woods (2006) noted that, if more than one habit is present, successful applications should focus on a single target behavior. In this case, a habit hierarchy is established based on the participant's subjective rating of the distress elicited by each habit (Piacentini & Chang, 2005), and each behavior is addressed sequentially from most to least distressing. Finally, the third session includes a review of the HRT procedure and the investigation of any problems that may have arisen throughout the course of treatment (Woods, 2006).

Primary Ingredients of HRT

The complex nature of the original conceptualization of HRT has led to attempts by a number of researchers to identify the active primary components of the procedure and to elucidate which elements were necessary for the effectiveness of the package. Many authors considered HRT to be a lengthy and overly comprehensive procedure (O'Connor, 2005a).

ACCEPTED MANUSCRIPT

Subsequently, researchers have evaluated a number of variations of HRT for the treatment of habit disorders (O'Connor, 2005a). The majority of component analysis studies addressed varying combinations of awareness training, competing response training, and social support (Miltenberger et al., 1998).

Allen (1998) employed awareness training and competing response training in a boy whose disruptive aggressive outbursts moderated within a five-month period. Miltenberger, Fuqua, and McKinley (1985) found that the awareness and competing response training components of HRT were as effective for muscle tics as the entire procedure specified by Azrin and Nunn (1973). Further investigation involving the sequential application of HRT components has shown that awareness training alone is not sufficient to reduce the manifestation of tics (O'Connor, 2005a). In a review, Wood and Miltenberger (1995) concluded that the essential elements of HRT are awareness training and competing response training. A study by Flessner et al. (2005) showed that HRT was just as effective for nail biting without the social support component as with it. However, research has shown that inclusion of the social support component is important for effective treatment of children and individuals with recalcitrant habits (O'Connor, 2005a; Woods & Miltenberger, 1995). Simplified versions of HRT consisting of awareness, competing, and social support training are commonly described in the literature as Simplified Habit Reversal Therapy (SHRT).

In additional variations, research addressing the necessity of an antagonistic competing response has found that the topography of the response need not be similar to the target behavior to produce reductions in tics (Sharenow, Fuqua, & Miltenberger, 1989; Verdellen, Keijsers, Cath, & Hoodguin, 2004). However, engaging in the competing response at the same time as the urge to perform the habitual behavior has been found necessary to decrease these behaviors (Miltenberger & Fuqua, 1985). Therefore, the competing response must be dependent upon the urge to perform the target behavior to ensure its effectiveness.

ACCEPTED MANUSCRIPT

Additionally, SHRT has been modified to optimally address stuttering behaviors (Azrin & Nunn, 1974). This procedure employs a modified competing response in which clients learn to regulate their breathing when speaking. The regulated-breathing method acts as an incompatible activity to stuttering behavior and has been successfully applied to both adults (Azrin & Nunn, 1974; Azrin, Nunn, & Frantz, 1979) and children (Elliot, Miltenberger, Rapp, Long, & McDonald, 1998).

Efficacy of HRT

Many studies have examined the efficacy of HRT for habit disorders. Clinical case studies involving tics have consistently reported target behavior reductions of 75% to 100% with HRT (O'Connor & Leclerc, 2008). Randomized controlled trials have also provided support for the efficacy of HRT with tics, starting with Azrin, Nunn and Frantz (1980a). Carr and Chong (2005) and Cook and Blacher (2007) reviewed the empirical evidence on use of HRT to ameliorate tics and found substantial evidence of efficacy. Less aggregation of results has occurred with regard to the efficacy of HRT with habit disorders other than tics. So far, no studies have quantified the effect size of HRT across studies and examined whether effect size varies with type of habit disorder or with other variables. The purpose of the present meta-analysis was to obtain an estimate of the efficacy of HRT and its variants, SHRT, and Regulated Breathing Method (RBM), for repetitive behavior disorders, and to examine possible moderators of effect size, such as disorder type and type of HRT.

Method

Literature Search

A systematic literature search included PsycINFO, Cambridge Scientific Abstracts, ProQuest, Google Scholar, and PubMed online databases for studies published prior to or during July 2009. We extracted studies containing "habit reversal" as a keyword and at least one of the following search terms: (a) cognitive-behavioral or behavioral, (b) intervention or

ACCEPTED MANUSCRIPT

treatment or therapy, (c) control group or placebo or comparison, (d) habit disorder, repetitive behavior disorder, Tourette's syndrome, or tics, and (e) trichotillomania, skin-picking, stuttering, nail-biting, or thumb-sucking.

To cover the research on HRT as exhaustively as possible and to prevent bias in the inclusion of studies, we used a series of literature search strategies. First, we checked all major narrative review articles and empirical studies with the potential for inclusion in the meta-analysis for citations to other relevant literature. Second, to minimize the "file-drawer effect," we contacted authors with multiple contributions to the habit reversal literature regarding the existence and availability of unpublished research. However, we obtained no usable unpublished studies by contacting authors.

Inclusion and exclusion criteria

In order to be included in the review and meta-analysis, studies had to meet the following inclusion criteria: (a) the study must have been designed to assess the efficacy of HRT for reducing tics, body-focused repetitive behaviors, or stuttering, (b) participants must have been randomly allocated to either HRT or a comparison condition, (c) the protocol for HRT must have included awareness training and competing behavior training, (d) the HRT condition must not have included any significant treatment procedure that is not part of HRT as originally described by Azrin and Nunn (1973), and (e) the study must have provided sufficient data to allow calculation of an effect size.

We excluded pre-post studies because of the lack of a control condition. We excluded studies using single-case designs because (a) such studies lack important design elements such as a control group, (b) the studies, with low *N*, would have little impact on a meta-analysis that included between-group studies, and (c) there is no current consensus regarding the calculation of effect sizes in single case studies.

ACCEPTED MANUSCRIPT

We excluded one HRT study (Teng, Woods, & Twohig, 2006, skin picking) from the meta-analysis because the report contained insufficient information to calculate an effect size We excluded the following studies because they mixed HRT with other significant treatment methods, such as self-punishment and cognitive therapy, in the experimental condition: Ehlers, Strangier, & Gieler, 1995; Ninan, Rothbaum, Marsteller, Knight, & Eccard, 2000; O'Connor, Brault, Robillard, Loiselle, Borgeat, & Stip, 2001 (tic and habit disorders); O'Connor, Gareau, & Borgeat, 1997; Van Minnen, Hoogduin, Keijsers, Hellenbrand, & Hendriks, 2003; Woods, Wetterneck, & Flessner, 2006.

Moderator Coding

We coded studies into commonly available variables that might be relevant to effect size. The variables coded into categories included client gender (female and mixed; there were no studies using only male clients), client age group (children, adult, both children and adult), disorder type (Tourette's syndrome; thumb-sucking; stuttering; nail-biting; mixed oral habits such as thumb sucking, nail biting and lip or tongue biting; temporomandibular disorder -- TMD), type of outcome measures (observed behavior rate, self-reported behavior frequency), therapy format (individual, group, parents of thumb suckers), comparison group type (bona fide other treatment, attention/placebo comparison, waiting list group), and whether the study included a developer of HRT, in this case Nathan Azrin, as one of the researchers (yes or no). Variables coded continuously included date of publication, total hours of therapist contact, study length (coded in months from pre-intervention assessment to final assessment that included both HRT and comparison participants), and methodological quality. For methodological quality, we considered study-design characteristics deemed important by Cuijpers, van Stratten, Schuurmans van Oppen, Holland, and Andersson (2010) and Gearing, El-Bassel, Ghesquierre, Baldwin, Gillies and Ngeow (2010) and calculated the total number of the following four design characteristics: (a) used treatment manual or

ACCEPTED MANUSCRIPT

written protocol, (b) HRT therapists specifically trained to provide HRT, (c) treatment integrity checked, e.g., by viewing videotapes, and (d) observation or interview schedule by assessor blinded as to condition or as to whether pre or post used as an outcome measure included in determination of effect size. We scored 1 for each characteristic clearly applied in the study and 0 when the characteristic was not applied or no mention was made of the characteristic in the article. Hence, method-quality scores could range from 0 to 4. We considered coding the studies for whether they included a diagnosis, but we found that all the Tourette's syndrome and TMD studies had diagnoses, and none of the other studies, which involved problems that did not have generally followed diagnostic criteria, did not have diagnoses. So we completed only an analysis of the different disorder types.

The meta-analysis had three coders who independently coded the articles. When discrepancies occurred, the coders discussed the matter until they reached consensus on the coding.

Statistical Analyses

Several studies provided outcome data for more than one follow-up period following treatment completion. In such cases, the longest follow-up period was used. Where follow-up data was not available for all conditions due to attrition within the conditions, the longest period at which 75% of the data was available was included in the meta-analysis. If no mention was made regarding the attrition rate of the sample, no drop-out rate was assumed.

Two studies presented other issues for coding. The study of Woods, Murray, Fuqua, Seif, Boyer, and Siah (1999) did not report standard deviation values for pre-treatment in both the HRT and control conditions. In this case, we used the HRT and control group post-treatment standard deviation values as an estimate for these values. In the study of Verdellen et al. (2004), all participants received both exposure therapy and HRT prior to follow-up assessment, so we did not use follow-up data from this study.

ACCEPTED MANUSCRIPT

Cohen's *d* (standardized mean-difference) effect size was used in the meta-analysis. For studies reporting group differences, effect sizes were calculated based on means and standard deviations and if these were not available on *t* test values. Effect sizes were estimated from percentage-decrease values for several studies. All computations were based on guidelines provided by Lipsey and Wilson (2001).

We used inverse variance weighting ($w = 1/SE$) macros following equations recommended by Lipsey and Wilson (2001) to compute descriptive and inferential statistics. Homogeneity analyses were conducted using the *Q* statistic. The effect sizes for the outcome variables across all studies were examined for univariate outliers (criterion $z = 3.30$, $p = .001$). No outliers were found. Moderator testing was conducted using meta-analysis statistics analogous to ANOVA, and continuous variables such as follow-up period were examined using statistics analogous to multiple regression analyses (Lipsey & Wilson, 2001).

Results

Eighteen studies produced effect sizes, with a total of 19 between-groups comparisons and a total of 575 participants. Table 1 summarizes key characteristics of the studies included in the meta-analysis, including sample characteristics and effect size. The overall effect of HRT was large (Cohen, 1988), $d = 0.80$ [95% *CI*: 0.53, 1.07], with 57 as the fail-safe *N* (Borenstein, Hedges, Higgins, & Rothstein, 2009), specifically the number of zero-effect-size studies needed to reduce *d* to 0.20, an effect size too small to have importance. The homogeneity analysis indicated a significant level of heterogeneity among effect sizes, $Q(18) = 37.42$, $p = .005$. The heterogeneity of effect sizes led us to adopt a random-effects model for the overall effect size and moderator analysis. This model generally produces larger effect sizes and larger confidence intervals than a fixed-effects model (Lipsey & Wilson, 2001).

ACCEPTED MANUSCRIPT

Table 2 shows the moderator analysis with nominal variables. No level of any moderator produced any less than a medium effect size (*d* = 0.50 or higher). No moderators were statistically significant. Other analyses explored three continuous potential moderators. None of the three was significantly related to effect sizes: date of publication, beta = -.21, *N* = 19, *p* = .35; number of hours of therapist contact, beta = -.36, *N* = 15, *p* = .14; follow-up time beta = .01, *N* = 17, *p* = .98; method quality score, beta = -.18, *N* = 19, *p* = .45. The mean number of months of follow-up time (time from pre-assessment to follow up assessment) was 4.34, with a range of 1 to 13 months. On average the studies had about one of the four high-quality method characteristics examined in the meta-analysis, which were use of a manual, specific training for therapists, assessment of fidelity of treatment sessions, and use of an assessor blinded to treatment condition. One study (Piancentini et al., 2010) had all four characteristics, and one study (Verdellen et al., 2004) had all of the characteristics except blind assessment.

Discussion

The present study found that the overall effect of HRT from pre-assessment to final post-treatment assessment was large in comparison with control conditions. The overall finding suggests that HRT is an efficacious behavioral intervention for the symptoms of disorders for which it is commonly used: tics; stuttering; and habit disorders such as nail biting, thumb sucking, and lip biting. The large effect size, *d* = 0.80, roughly equivalent across disorders, is noteworthy. Furthermore, the moderator analysis showed HRT produces stable and enduring treatment effects through reductions in maladaptive habitual behavior that extend beyond the period of intervention. These findings are consistent with the findings of a between-group study of skin picking that lacked essential information for a meta-analysis (Teng et al., 2006) and uncontrolled case-studies and systematic small-N studies investigating the effectiveness of HRT that did not meet the methodological standards for inclusion in the

ACCEPTED MANUSCRIPT

meta-analysis (see, e.g., O'Connor & Leclerk, 2008; Miltenberger, Wagaman, & Arndorfer, 1996). The results are also consistent with the conclusions of the literature review of Keuthen et al. (2005). The current findings add meta-analytic effect sizes and moderator evaluation to previously completed studies and reviews.

Seventeen of the 19 treatment analyses in the meta-analysis showed superior effects for HRT over a comparison group. The two exceptions were the studies of Verdellen et al. (2004) and Glaros et al. (2007). Verdellen et al. (2004) reported a greater effect for exposure with response prevention (ERP) than for HRT. However, as noted in Table 1, the comparison treatment involved twice as much time in therapy as for HRT, arguably creating a confound in favor of ERP. In the Glaros et al. study the comparison treatment for temporomandubular disorder, a splint, had a better outcome than the HRT group, but with a total sample size of eight the results have a large confidence interval (95% CI: -1.77, 1.02), suggesting the need for a replication with a larger sample size.

The effect sizes for HRT were similar across specific type of disorder, suggesting a treatment-relevant commonality in tic disorders, habit disorders, and stuttering. This result indicates that HRT tends to have positive effects on the types of disorders on which it is commonly used.

Effect size was non-significantly lower with regard to bona fide comparison conditions than with regard to waiting list and attention/placebo conditions. The moderate effect size of HRT, $d = 0.73$, in comparison to bona fide treatments, suggests that HRT is at present the treatment of choice for tics, habit disorders, and stuttering.

Cook and Blacher (2007) concluded that HRT satisfied the criteria for classification as a well-established treatment for tic disorders. The findings of the present meta-analysis are consistent with that conclusion and indicate that HRT is also a well-established treatment for repetitive-behavior disorders that involve picking, plucking, biting, and scratching. The

ACCEPTED MANUSCRIPT

results regarding stuttering are consistent with the conclusion that HRT is probably efficacious for stuttering (Conelea et al., 2006). Across all the types of habit disorders studies, tic-related and other, the studies meet the standards of a well-established treatment in that they provide evidence of efficacy in comparison to other bona fide treatments or attention-placebo conditions, with adequate study sample sizes, replication by different researchers, and use of detailed or manualized treatment methods (Cook & Blacher, 2007). This conclusion, of course, assumes that different habit disorders have enough in common to warrant a treatment-efficacy conclusion about the entire group. The evidence is clearest with regard to stuttering, thumb sucking, nail biting, and TMJ, which each have multiple studies showing the efficacy of HRT.

The number of hours of therapist contact was not related to treatment effect size. In fact, the nonsignificant trend was in the opposite direction. Individual studies involving no clinician contact produced a comparable overall effect size to studies employing therapist interaction. This finding is important if one considers the costs of therapist time (Himle, Flessner, Bonow, & Woods, 2006). The finding is consistent with efficacy studies of various types of interventions with low therapist involvement for diverse disorders (see, e.g., Barak, Hen, Boniel-Nissim, & Shapira, 2008; Vocks, Tuschen-Caffier, Pietrowsky, Rustenbach, Kersting, & Herpertz, 2010).

One other moderator narrowly missed significance – whether one of the researchers was a developer of HRT. In this case, the researcher was Nathan Azrin. With regard to problem-solving therapy, one of the developers had significantly higher effect sizes than other researchers (Malouff, Thorsteinsson, & Schutte, 2007), so this variable is worthy of consideration when evaluating the effects of interventions developed by one of the researchers.

ACCEPTED MANUSCRIPT

Study methodological characteristics that make results more convincing, including use of a treatment manual, training of therapists specifically for the treatment, assessment of fidelity of delivery of the treatment, and use of an assessor blinded to treatment condition, were not significantly associated with effect size. It appears that using these methods has no impact on the effects of HRT, i.e., the lower quality methods produce about the same effect sizes as better quality studies. Consistent with this view is the finding that type of comparison group and time from pre-assessment to final assessment were not significantly related to effect size.

Regarding outcome measures, the overall effect size for HRT treatment effect did not vary significantly according to assessment method. This finding suggests that the use of subjective measurement approaches to quantify the severity and frequency of target behavior does not lead to greater effect sizes in studies of habit disorders. However, self-report and direct observation measures together may provide the most comprehensive and accurate account for the symptom frequency, intensity, complexity, and duration, as well as information concerning the probability of psychosocial impact (Himle et al., 2006; Keuthen et al., 2005).

The effect size of HRT did not decrease significantly with longer study times. In fact, the non-significant trend was in the opposite direction. This important finding suggests continuing positive effects over time, at least over a range of up to 13 months from pre-intervention.

Despite the effectiveness of HRT over other treatments for habit disorders, HRT has limits. Not all clients respond to HRT (Keuthen et al., 2005), and some clinicians view HRT as an incomplete intervention as it fails to address private events (e.g., emotions; Diefenbach et al., 2002). Miltenberger (2008, p. 489) recently suggested alternative or supplemental treatment methods, such as differential reinforcement of other behavior and response cost, for

ACCEPTED MANUSCRIPT

cases in which HRT is not effective enough. For tics, exposure with response prevention might also serve as an alternative treatment (see Verdellan, 2004).

Some of the published studies evaluating HRT have lacked important information, such as the length of treatment sessions and the length of the follow-up. Almost all of them lacked a specified effect size, such as Cohen's *d*, for the main analyses. Cohen's *d* is valuable because it expresses the between-group mean difference in outcome over the pooled standard deviation of the sample. Hence, a *d* of 1.0 means that on average clients in one condition finished the study one standard deviation higher than clients in the other condition.

The method of the current meta-analysis has limitations that warrant consideration. First, different moderators sometimes overlap with each other with regard to specific studies, making it difficult to determine which moderator has a specific effect. Second, the moderator analyses were not based on random assignment to a moderator level. Hence, the results of the moderator analyses are merely suggestive. Third, only a small number of studies examined some levels of moderators, such as nail biting as the problem. Thus, not all levels of some moderators received adequate evaluation.

The strengths of the meta-analysis include the inclusion of hundreds of participants across studies by various researchers using somewhat different HRT versions, measures, and therapists, applied to various habit disorders. By combining the results of many studies statistically, the meta-analysis provides a quantification of effect size with examination of whether the effects vary with conditions of the treatment or the clients.

Future research on HRT might most profitably examine its efficacy when delivered with minimal or no therapist contact. To the extent that such treatment is efficacious, it has importance in its low cost and potentially wide application. Also, it could be useful to examine the causes of treatment failure, when that occurs. The more high-quality research design elements incorporated in the research and mentioned in the report, the more

ACCEPTED MANUSCRIPT

convincing the results will be. Research findings will be most useful if the information provided in the report includes the length of treatment sessions, information about the format of therapy (e.g., individual vs. group), a clear statement of the timing of the follow-up, information about whether therapists had specific training for the therapy, whether they had a treatment manual, to what extent they followed the treatment manual in important respects, and a specific between-groups effect size, preferably Cohen's *d*.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

References

References marked with an asterisk indicate studies included in the meta-analysis.

Adams, A. N., Adams, M. A., & Miltenberger, R. G. (2008). Habit reversal training. In W. O'Donohue & J. E. Fisher (Eds.). *Cognitive behavior therapy: Applying empirically supported techniques in your practice.* Hoboken, NJ: Wiley..

Allen, K. A. (1998). The use of an enhanced simplified habit-reversal procedure to reduce disruptive outbursts during athletic performance. *Journal of Applied Behavior Analysis, 31*, 489-492.

Azrin, N. H., & Nunn, R. G. (1973). Habit-Reversal: A method of eliminating nervous habits and tics. *Behaviour Research & Therapy, 11*, 141-149.

Azrin, N. H., & Nunn, R. G. (1974). A rapid method of eliminating stuttering by a regulated breathing approach. *Behaviour Research & Therapy, 12*, 279-286.

Azrin, N. H., & Nunn, R. G. (1977). *Habit control in a day.* Simon and Schuster: New York.

Azrin, N. H., Nunn, R. G., & Frantz, S. E. (1979). Comparison of regulated-breathing versus abbreviated desensitization on reported stuttering episodes. *Journal of Speech and Hearing Disorders, 44*(3), 331-339.

Azrin, N. H., Nunn, R. G., & Frantz, S. E. (1980a). Habit reversal vs negative practice treatment of nervous tics. *Behavior Therapy, 11*, 169-178.

*Azrin, N. H., Nunn, R. G., & Frantz, S. E. (1980b). Treatment of hair-pulling (Trichotillomania): A comparative study of habit reversal and negative practice training. *Journal of Behavior Therapy and Experimental Psychiatry, 11*, 13-20.

ACCEPTED MANUSCRIPT

*Azrin, N. H., Nunn, R. G., & Frantz, S. E. (1980c). Habit reversal vs. negative practice treatment of nailbiting. *Behaviour Research and Therapy, 18*, 281-285.

*Azrin, N. H., Nunn, R. G., & Frantz-Renshaw, S. E. (1982). Habit reversal vs. negative practice treatment of self-destructive oral habits (self-biting, chewing or licking of the lips, cheeks, tongue or palate). *Journal of Behavior Therapy and Experimental Psychiatry, 13*, 49-54.

*Azrin, N. H., Nunn, R. G., & Frantz-Renshaw, S. E. (1980). Habit reversal treatment of thumbsucking. *Behavior Research & Therapy, 18*, 395-399.

*Azrin, N. H., & Peterson, A. L. (1990). Treatment of Tourette's Syndrome by habit reversal: A waiting-list control group comparison. *Behavior Therapy, 21*, 305-318.

Barak, A., Hen, L., Boniel-Nissim, M., & Shapira, N. (2008). A comprehensive review and a meta-analysis of the effectiveness of Internet-based psychotherapeutic interventions. *Journal of Technology in Human Services, 26*, 109-120.

Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. New York: Wiley.

Carr, J. E., & Chong, I. M. (2005). Habit reversal treatment of tic disorders: A methodological critique of the literature. *Behavior Modification, 33*, 455-456.

*Christensen, A. P., & Sanders, M. R. (1987). Habit reversal and differential reinforcement of other behavior in the treatment of thumb-sucking: An analysis of generalization and side-effects. *Journal of Child Psychology and Psychiatry, 28*(2), 281-295.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.

ACCEPTED MANUSCRIPT

Conelea, C. A., Rice, K. A., & Woods, D. W. (2006). Regulated breathing as a treatment for stuttering: a review of the evidence. Regulated breathing as a treatment for stuttering: A review of the empirical evidence. *Journal of Speech and Language Pathology and Applied Behavior Analysis*, *1*, 94-102

Cook, C. R., & Blacher, J. (2007). Evidence-based psychosocial treatments for tic disorders. *Clinical Psychology: Science and Practice, 14*, 252-267.

Cuijpers, P., van Stratten, A., Schuurmans, J., van Oppen, P., Holland, S. D., & Andersson, G. (2010). Psychotherapy for chronic major depression and dysthymia: A meta-analysis. *Clinical Psychology Review, 30*, 51-62.

*Deckersbach, T., Rauch, S., Buhlmann, U., & Wilhelm, S. (2006). Habit reversal versus supportive psychotherapy in Tourette's disorder: A randomised controlled trial and predictors of treatment response. *Behaviour Research and Therapy, 44,* 1079-1090.

Diefenbach, G. J., Reitman, D., & Willianson, D. A. (2000). Trichotillomania: A challenge to research and practice. *Clinical Psychology Review, 29,* 289-309.

Ehlers, A., Strangier, U., & Geiler, U. (1995). Treatment of atopic dermatitis: A comparison of psychological and dermatological approaches to relapse prevention. *Journal of Consulting and Clinical Psychology*, *63*, 624-635.

Elliot, A. J., Miltenberger, R. G., Rapp, J., Long, E. S., & McDonald, R. (1998). Brief application of simplified habit reversal to treat stuttering in children. *Journal of Behavior Therapy and Experimental Psychiatry, 29,* 289-302.

Flessner, C. A., Miltenberger, R. G., Egemo, K., Kelso, P., Jostad, C., Johnson, B. et al. (2005). An evaluation of the social support element of simplified habit reversal. *Behavior Therapy, 36,* 35-42.

ACCEPTED MANUSCRIPT

Gearing, R. E., El-Bassel, N., Ghesquiere, A., Baldwin, S., Gillies, J., & Ngeow, E. (2010). Major ingredients of fidelity: A review and scientific guide to improving quality of intervention research implementation. *Clinical Psychology Review, 31,* 79-88.

*Glaros, A. G., Weroha, N. K., Lausten, L., & Franklin, K. L. (2007). Comparison of habit reversal and a behaviorally-modified dental treatment for temporomandibular disorders. *Applied Psychophysiology and Biofeedback, 32*, 149-154.

Himle, M. B., Woods, D. W., Conelea, C. A., Bauer, C. C., & Rice, K. A. (2007). Investigating the effects of tic suppression on premonitory urge ratings in children and adolescents with Tourette's Syndrome. *Behaviour Research & Therapy, 45*, 2964-2976.

*Horne, D. J., & Wilkenson, J. (1980). Habit reversal treatment for fingernail biting. *Behaviour Research & Therapy, 18*, 287-291.

Keuthen, N. J., Bohne, A., Himle, M., & Woods, D. W. (2005). Advances in the conceptualization and treatment of body-focused repetitive behaviors. In B. E. Ling (Ed.), *Obsessive compulsive disorder research* (pp. 1-29). Hauppauge, NY: Nova Biomedical Books.

*Ladouceur, R., & Martineau, G. (1982). Evaluation of a regulated-breathing method with and without parental assistance in the treatment of child stutterers. *Journal of Behavior Therapy and Experimental Psychiatry, 13*, 301-306.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.

Malouff, J. M., Thorsteinsson, E., & Schutte, N. S. (2007). The efficacy of problem solving therapy in reducing mental and physical health problems: A meta-analysis. *Clinical Psychology Review*, 27, 46-57.

ACCEPTED MANUSCRIPT

Miltenberger, R. G. (2008). *Behavior modification: Principles and procedures* (4th ed.). Thomson: Belmont, CA.

Miltenberger, R. G., & Fuqua, R. W. (1985). A comparison of contingent vs non-contingent competing response practice in the treatment of nervous habits. *Journal of Behavior Therapy and Experimental Psychiatry, 16*, 195-200.

Miltenberger, R. G., Fuqua, R. W., & McKinley, T. (1985). Habit reversal with muscle tics: Replication and component analysis. *Behavior Therapy, 16*, 39-50.

Miltenberger, R. G., Fuqua, R. W., & Woods, D. W. (1998). Applying behavior analysis to clinical problems: Review and analysis of habit reversal. *Journal of Applied Behavior Analysis, 31*, 447-469.

Miltenberger, R. G., Wagaman, J. R., & Arndorfer, R. E. (1996). Simplified treatment and long term follow-up for stuttering in adults: A study of two cases. *Journal of Behaviour Therapy & Experimental Research, 27*, 181-188.

Ninan, P. T., Rothbaum, B. O., Marsteller, F. A., Knight, B. T., & Eccard, M. B. (2000). A placebo-controlled trial of cognitive behavioral therapy and clomipramine in trichotillomania. *Journal of Clinical Psychiatry, 61*, 47-50.

O'Connor, K. P. (2005a) *Cognitive-behavioral management of tic disorders*. Chichester, UK: Wiley.

O'Connor, K. P. (2005b). A cognitive-behavioral/psychophysiological model of tic disorders. *Behaviour Research and Therapy, 40*, 1113-1142.

O'Connor, K. P., Brault, M., Robillard, S., Loiselle, J., Borgeat, F., & Stip, E. (2001). Evaluation of a cognitive-behavioural program for the management of chronic tic and habit disorders. *Behaviour Research & Therapy, 39*, 667-681.

ACCEPTED MANUSCRIPT

O'Connor, K., Gareau, D., & Borgeat, F. (1997). A comparison of a behavioral and a cognitive-behavioral approach to the management of chronic tic disorders. *Clinical Psychology and Psychotherapy, 4*, 105-117.

O'Connor, K., & Leclerc, J. (2008). Tic disorders and Tourette Syndrome. In J. S. Abramowitz, D. McKay, & S. Taylor (Eds.), *Clinician's handbook of obsessive-compulsive disorder and related problems* (pp. 270-287). Baltimore, MD: Johns Hopkins University Press.

Piacentini, J., & Chang, S. (2005). Habit reversal training for tic disorders in children and adolescents. *Behavior Modification, 29*, 803-821.

Piacentini, J., Woods, D. W., Scahill, L., Wilhelm, S., Peterson, A. L., Chang, S. et al. (2010). Behavior therapy for children with Tourette Syndrome: A randomized controlled trial. *JAMA, 303*, 1929-1937.

Rosenbaum, M. S., & Ayllon, T. (1981). The behavioral treatment of neurodermatitis through habit-reversal. *Behaviour Research & Therapy, 19*, 313-318.

*Saint-Laurent, L., & Ladouceur, R. (1987). Massed versus distributed application of the regulated-breathing method for stutterers and its long-term effect. *Behavior Therapy, 18*, 38-50.

Sharenow, E. L., Fuqua, R. W., & Miltenberger, R. G. (1989). The treatment of muscle tics with dissimilar competing response practice. *Journal of Applied Behavior Analysis, 22*, 35-42.

Teng, E. J., Woods, D. W., & Twohig, M. P. (2006). Habit reversal as a treatment for chronic skin picking: A pilot investigation. *Behavior Modification, 30*, 411-422.

ACCEPTED MANUSCRIPT

*Townsend, D., Nicholson, R. A., Buenaver, L., Bush, F., & Gramling, S. (2001). Use of a habit reversal treatment for temporomandibular pain in a minimal therapist contact format. *Journal of Behavior Therapy and Experimental Psychiatry, 32*, 221-239.

Twohig, M. P., & Woods, D. W. (2001). Habit reversal as a treatment for chronic skin picking in typically developing adult male siblings. *Journal of Applied Behavior Analysis, 34*, 217-220.

*Twohig, M. P., Woods, D. W., Maracks, B. A., & Teng, E. J. (2002). Evaluating the efficacy of habit reversal: comparison with a placebo control. *Journal of Clinical Psychiatry, 64*, 40-48.

Van Minnen, A., Hoogduin, K. A. L., Keijsers, G. P. J., Hellenbrand, I., & Hendricks, G. L. (2003). Treatment of trichotillomania with behavioral therapy or fluoxetine. *Archives of General Psychiatry, 60*, 517-522.

*Verdellen, C. W. J., Keijsers, G. P. J., Cath, D. C., & Hoodguin, C. A. L. (2004). Exposure with response prevention versus habit reversal in Tourette's Syndrome: A controlled study. *Behaviour Research and Therapy, 42*, 501-511.

Vocks, S., Tuschen-Caffier, B., Pietrowsky, R., Rustenbach, S. J.; Kersting, A., & Herpertz, S. (2010). Meta-analysis of the effectiveness of psychological and pharmacological treatments for binge eating disorder. *International Journal of Eating Disorders, 43*, 205-217

*Waterloo, K. K., & Gotestam, K. G. (1988). The regulated-breathing method for stuttering: An experimental evaluation. *Journal of Behavior Therapy and Experimental Psychiatry, 19*, 11-19.

ACCEPTED MANUSCRIPT

*Wilhelm, S., Deckersbach, T., Coffey, B. J., Pterson, A. L. & Baer, L. (2003). Habit reversal versus supportive psychotherapy for Tourette's Disorder: a randomized controlled trial. *American Journal of Psychiatry, 160*, 1175-1177.

Woods, D. W. (2006). Habit reversal treatment manual for tic disorders. In D. W. Woods & R. G. Miltenberger (Eds.), *Tic disorders, trichotillomania, and other repetitive behavior disorders* (pp. 97-132). US: Springer.

Woods, D. W., & Miltenberger, R. G. (1995). Habit reversal: A review of the applications and variations. *Journal of Behavior Therapy and Experimental Psychiatry, 26*(2), 123-131.

*Woods, D. W., Murray, L. K., Fuqua, R. W., Seif, T., Boyer, L. J., & Siah, A. (1999). Comparing the effectiveness of similar and dissimilar competing responses in evaluating the habit reversal treatment for oral-digital habits in children. *Journal of Behavior Therapy and Experimental Psychiatry, 30*, 289-300.

Woods, D. W., Wetterneck, C. T., & Flessner, C. A. (2006). A controlled evaluation of acceptance and commitment therapy plus habit reversal for trichotillomania: *Behaviour Research & Therapy, 44*, 639-656.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Table 1
Key Descriptive Characteristics of Studies in the Meta-Analysis

Study	*N*	Gender	Age group	Disorder	Therapy format	Treatment time (minutes)	Study period (months)	Comparison	*d*	*SE*
Azrin,	3	mix	Both	Trichotillom	Indivi	120	3.00	Massed prac	1.0	0.3
Azrin,	9	mix	Adult	Nail biting	Grou	120	3.00	Massed prac	1.1	0.2
Azrin,	1	mix	Both	Oral habits	Indivi	120	1.00	Massed prac	1.2	0.6
Azrin.	2	mix	Childr	Thumb	Indivi	120	3.00	Bitter subst	1.2	0.4
Azrin &	1	mix	Both	Tourette's	Indivi	-	-	Waiting list	1.9	0.7
Christe	2	mix	Childr	Thumb	Paren	120	3.00	DRO	0.2	0.4
Christe	2	mix	Childr	Thumb	Paren	120	3.00	Waiting list	0.9	0.4
Deckers	3	mix	Adult	Tourette's	Indivi	700	-	SP	1.6	0.4
Glaros,	8	fem	Adult	TMD	Indivi	-	13.0	Splint	-	0.7
Horne	1	mix	Adult	Nail biting	Indivi	-	3.50	Nail care	0.7	0.5
Ladouc	1	mix	Childr	Stuttering	Indivi	135	1.75	Waiting list	0.1	0.5
Piacentini, et al. (2010)	12 6	mixed	Children[c]	Tourette's	Individual	540	2.50	Attention placebo treatment	0.68	0.18
Saint-	4	mix	Adult	Stuttering	Grou	1,920	2.00	Self-record	0.1	0.4
Townse	9	fem	Adult	TMD	Indivi	318	5.00	Waiting list	1.1	0.5
Twohig	1	mix	Adult	Nail biting	Indivi	360	5.75	Discussion	1.2	0.5
Verdell	3	mix	Both	Tourette's	Indivi	600	6.25	ERP[g]	-	0.3
Waterlo	3	mix	Adult	Stuttering	Indivi	150	9.00	Waiting list	1.8	0.4
Wilhel	2	mix	Adult	Tourette's	Indivi	-	7.50	SP	0.6	0.4
Woods	1	mix	Childr	Oral-digital	Indivi	360	1.60	Waiting list	0.1	0.4

Note. HRT type: HRT=original version; SHRT=simplified version; RBM=Regulated Breathing Method. Disorder. Disorder Type: TMD = Temporomandibular Disorder (grinding/chewing). Comparison: Massed prac = massed practice; Bitter subst = foul-tasting substance on nails; DRO = differential reinforcement of any behavior other than thumb sucking; SP = supportive psychotherapy; self-record = self-recording of stuttering; ERP = exposure with response prevention. Oral habits = biting, chewing or licking lips, mouth, or tongue. Oral-digital = nail biting or thumb sucking.
[a] Treatment time does not include phone calls of unspecified length.
[b] Follow up results not used for analyses b/c los rate > 30%.
[c] Participants also included some adolescents; age group set as children.
[d] Effect size based on the mean across four HRT treatment conditions; Figure 1 p.44.
[e] Effect size is based on the mean pain variable because it was the most meaningful.
[f] Number of participants at follow up estimated based on information in the article.
[g] ERP clients had twice as much therapy time as HRT clients, 1200 minutes to 600.
[h] Effect size based on the mean score across three situations.
[i] Effect size based on the mean impairment and tic severity.

ACCEPTED MANUSCRIPT

Table 2
Overall Effect Size, Homogeneity Analysis for Moderator Analysis: Pre-Assessment to Follow-up Assessment for Habit Reversal Therapy (HRT)

			CI95%			Homogeneity Analysis		
Moderator	*n*	*d*	Lower	Upper	*p*	*Q*	*df*	*p*
Gender, $Q(1) < 0.23$, $p = .63$								
Female	2	0.57	-0.44	1.57	.273	2.07	1	.15
Mixed	17	0.82	0.53	1.11	<.001	16.52	16	.42
Age group, $Q(2) = 1.55$, $p = .46$								
Children	6	0.59	0.12	1.06	.014	2.57	5	.77
Adults	9	0.98	0.58	1.38	<.001	8.80	8	.36
Both (children and adults)	4	0.73	0.09	1.38	.026	6.05	3	.11
Disorder type, $Q(5) = 0.78$, $p = .978$								
Tourette's syndrome	5	0.78	0.16	1.39	.013	5.65	4	.23
Thumb-sucking	3	0.79	-0.02	1.60	.057	0.98	2	.61
Stuttering	3	0.74	-0.07	1.56	.074	3.81	2	.15
Nail-biting	3	1.06	0.27	1.85	.008	0.21	2	.90
Other oral habits	2	0.59	-0.52	1.71	.294	0.90	1	.34
TMD	2	0.53	-0.60	1.66	.359	1.67	1	.20
Therapy format, $Q(2) = 0.43$, $p = .81$								
Individual	15	0.85	0.52	1.19	<.001	14.82	14	.39
Group	2	0.70	-0.08	1.48	.078	1.40	1	.24
Parents of thumb suckers (group)	2	0.56	-0.35	1.47	.231	0.51	1	.47
Outcome measure, $Q(1) = 0.03$, $p = .86$								
Observed behavior rate or Physiological measure	12	0.78	0.40	1.16	<.001	11.70	11	.39
Self-report behavior frequency	7	0.83	0.41	1.25	<.001	6.55	6	.36
Comparison group type, $Q(1) = 0.62$ $p = .43$								
Waitlist group	6	0.98	0.45	1.52	<.001	6.55	5	.26
Other treatment	13	0.73	0.41	1.06	<.001	11.84	12	.46
Treatment developer as researcher, $Q(1) = 3.36$, $p = .07$								
No	14	0.66	0.35	0.96	<.001	17.11	13	.19
Yes	5	1.20	0.70	1.71	<.001	0.89	4	.93

Note. CI = confidence interval; Disorder Type: TMD = Temporomandibular Disorder (grinding/chewing).

ACCEPTED MANUSCRIPT

Highlights

- Habit reversal therapy has a large effect size in the treatment of tic disorders, habit disorders, and stuttering
- Habit reversal therapy tends to be efficacious for both adults and children
- The positive effects of habit reversal therapy tend to endure over time

ACCEPTED MANUSCRIPT