American Psychological Association

Journal of Consulting and Clinical Psychology

0022-006X/19/$12.00

2019, Vol. 87, No. 7, 590–602
http://dx.doi.org/10.1037/ccp0000405

Efficacy and Mediators of a Group Cognitive–Behavioral Therapy for Hoarding Disorder: A Randomized Trial

David F. Tolin
Institute of Living, Hartford, Connecticut, and Yale University School of Medicine

Bethany M. Wootton, Hannah C. Levy, Lauren S. Hallion, and Blaise L. Worden
Institute of Living, Hartford, Connecticut

Gretchen J. Diefenbach
Institute of Living, Hartford, Connecticut, and Yale University School of Medicine

James Jaccard
New York University

Michael C. Stevens
Institute of Living, Hartford, Connecticut, and Yale University School of Medicine

Objective: Hoarding disorder (HD) is a common and potentially debilitating psychiatric disorder. Thus far, psychological treatments have yielded modest effects and/or were time-consuming and costly to deliver. The aim of the present study was to test the efficacy of a brief group cognitive–behavioral therapy (CBT) for adults with HD and to test hypothesized mediators of treatment outcome. ***Method:*** Eighty-seven adults with a primary diagnosis of HD were randomized to either immediate CBT or wait list. CBT consisted of 16 weekly, 90-min group sessions that emphasized in-session practice of discarding and refraining from acquiring, decision-making and problem-solving training, emotional distress tolerance, motivational interviewing strategies, and contingency management. Participants were assessed at pretreatment, midtreatment, and posttreatment by an independent evaluator unaware of treatment condition. ***Results:*** CBT was efficacious for the symptoms of HD compared with wait list. Saving-related cognitions, but not subjective cognitive impairment, partially mediated treatment outcomes. ***Conclusion:*** Brief Group CBT is an efficacious and feasible treatment for adults with HD, and is partially mediated by reductions in maladaptive beliefs about possessions. Superiority trials comparing CBT to active treatments, and additional research into mechanisms of treatment outcome, are warranted.

What is the public health significance of this article?
Hoarding disorder (HD) is a common and often severe psychiatric illness. As a new diagnosis, it has received less research attention than many other disorders. Treatment efficacy for HD also lags behind that of many other psychiatric disorders. This study suggests that a relatively brief (16-session) Group CBT is efficacious for adults with HD, with 83% of patients classified as improved. These results support the consideration of Group CBT as a first-line treatment for HD.

Keywords: group therapy, obsessive–compulsive disorder, contingency management, behavior therapy

Supplemental materials: http://dx.doi.org/10.1037/ccp0000405.supp

This article was published Online First April 22, 2019.

David F. Tolin, Institute of Living, Hartford, Connecticut, and Department of Psychiatry, Yale University School of Medicine; Bethany M. Wootton, Hannah C. Levy, Lauren S. Hallion, and Blaise L. Worden, Institute of Living; Gretchen J. Diefenbach, Institute of Living, and Department of Psychiatry, Yale University School of Medicine; James Jaccard, Silver School of Social Work, New York University; Michael C. Stevens, Institute of Living, and Department of Psychiatry, Yale University School of Medicine.

Bethany M. Wootton is now at the Department of Psychology, University of Technology Sydney. Lauren S. Hallion is now at the Department of Psychology, University of Pittsburgh.

This work was funded by National Institute of Mental Health grant R01 MH101163. Clinicaltrials.gov identifier NCT01956344. The authors thank Amber Billingsley, Akanksha Das, Marla Genova, Benjamin Katz, and James Ransom for their assistance with data collection. The authors also thank Elizabeth Davis, Christina Gilliam, Scott Hannan, and Kristen Springer for serving as independent evaluators.

Correspondence concerning this article should be addressed to David F. Tolin, Institute of Living, 200 Retreat Avenue, Hartford, CT 06106. E-mail: david.tolin@hhchealth.org

Hoarding disorder (HD) was established as an independent diagnosis in the fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (*DSM–5*; American Psychiatric Association, 2013). The hallmark characteristic of HD is difficulty discarding possessions because of a perceived need to save them or significant distress upon discarding. The excessive saving behavior results in the accumulation of clutter that precludes the use of living spaces for their intended purpose (e.g., an individual with HD might no longer be able to cook in the kitchen or sleep in the bedroom because of clutter). Excessive acquiring behavior (either through buying or the accumulation of free objects) is listed as a specifier in *DSM–5*, though research suggests that acquiring behavior is present in most cases of HD (Frost, Rosenfield, Steketee, & Tolin, 2013), and current models of the disorder posit a critical role of acquiring behavior in illness maintenance (e.g., Kyrios, 2014).

Estimates of the prevalence of HD suggest that it is a common mental disorder, with prevalence rates in epidemiological studies ranging from 2%–4% (Mueller, Mitchell, Crosby, Glaesmer, & de Zwaan, 2009; Nordsletten et al., 2013). HD is associated with substantial functional impairment and poor health-related quality of life (Drury, Ajmi, Fernandez de la Cruz, Nordsletten, & Mataix-Cols, 2014), as well as significant public health cost (Tolin, Frost, Steketee, Gray, & Fitch, 2008). Current cognitive–behavioral models posit that individuals with HD save items because of impaired decision-making processes, maladaptive beliefs about possessions, and behavioral avoidance. Decision-making impairments may relate, in part, to impaired cognitive function as evidenced by neuropsychological testing (see Woody, Kellman-McFarlane, & Welsted, 2014) and standardized self-report (Samuels et al., 2018). Clinical observation suggests that patients with HD frequently seem to have difficulty focusing on topics and tasks and often become easily overwhelmed under task demands. Although evidence for executive function impairment is mixed, some evidence points to problems of categorization (e.g., Grisham, Norberg, Williams, Certoma, & Kadib, 2010) that comport well with behavioral observations of difficulty making rapid decisions, prioritizing, and distinguishing important from unimportant information. Maladaptive beliefs about possessions include a heightened sense of responsibility for possessions, excessive sentimental attachment to objects, lack of memory confidence, and exaggerated need for control (Steketee, Frost, & Kyrios, 2003). As a result of these processes, the act of decision-making is perceived as aversive, leading people with HD to avoid these activities, which ultimately leads to increased clutter.

Cognitive–behavioral therapy (CBT) targets these purported maintenance factors via efforts to improve decision-making, challenge maladaptive beliefs, and decrease behavioral avoidance (Steketee & Frost, 2014). The efficacy of individual CBT was determined in a randomized controlled trial (RCT), in which CBT proved superior to a wait-list control condition, with a high rate of clinical response (Steketee, Frost, Tolin, Rasmussen, & Brown, 2010). Across studies, the pretreatment to posttreatment effect size of CBT for HD is large ($g = 0.82$; Tolin, Frost, Steketee, & Muroff, 2015). However, it is also noted that the rate of clinically significant change (CSC), defined as scores closer to the normative range than to the HD range (Jacobson & Truax, 1991), is only 35%, indicating that most patients with HD remain significantly symptomatic after treatment (Tolin et al., 2015).

There is significant room for improvement of CBT for HD, from both a feasibility and an efficacy perspective. The protocol used in the initial trial of individual CBT is lengthy (twenty-six 120-min sessions) and includes home visits that are not feasible for most community clinicians. In the current study, the protocol was shortened to sixteen 90-min sessions, and home visits were not included to maximize disseminability (i.e., acceptability and feasibility for practicing clinicians). Treatment was also conducted in a group, rather than individual format, with the aim of improving feasibility and decreasing sense of stigma (Chasson, Guy, Bates, & Corrigan, 2018). Group CBT for HD appears promising (Bodryzlova, Audet, Bergeron, & O'Connor, 2018; Gilliam et al., 2011; Muroff, Steketee, Bratiotis, & Ross, 2012; Muroff et al., 2009), and pre- to posttreatment effects of individual versus Group CBT are comparable (Tolin et al., 2015). In the present study, several modifications to the CBT protocol were made in hopes of increasing treatment efficacy. First, the protocol (Table 1) includes regular in-session discarding practice (Steketee & Frost, 2014), in which patients bring possessions to the clinic and practice sorting and discarding. Second, problem-solving training is employed to improve cognitive efficiency. Third, emotional distress tolerance is emphasized as a means of emotion regulation, given findings that poor distress tolerance is associated with saving behavior (e.g., Shaw & Timpano, 2016). Fourth, strategies from motivational interviewing are included midway through treatment to improve compliance, which can be quite variable and is associated with treatment outcomes (Tolin, Frost, & Steketee, 2007). Finally, some elements of contingency management were also included to increase compliance; similar contingencies have shown promise when added to Group CBT in an open trial for HD (Worden, Bowe, & Tolin, 2017). An initial open trial ($N = 20$) of the present treatment protocol, combined with an Internet support program, showed large reductions in HD severity ($d = 1.57$) that were well above the average pre- to posttreatment effect in prior research (Ivanov et al., 2018).

In addition to examining the efficacy of Group CBT for HD, it is also important to understand mediators of clinical improvement, which are a necessary (though not sufficient) criterion for identifying mechanisms (see Tryon, 2018). Here, we examined two strong candidates: maladaptive cognitions and subjective cognitive impairment. Maladaptive cognitions such as a heightened sense of responsibility, excessive sentimental attachment, poor memory confidence, and need for control have been proposed to play a central mechanistic role in HD (Steketee et al., 2003). A previous RCT that utilized a somewhat more cognitively based individual treatment for HD (Steketee et al., 2010) found that reductions in these maladaptive beliefs mediated HD symptom reduction among treatment completers (Levy et al., 2017). It remains to be seen whether cognitive change would also mediate outcomes in the present group-based treatment, which is more skill-based than cognitively focused, or whether such an intervention would operate via an alternate mechanism. However, previous research in related disorders such as depression suggests that changes in maladaptive cognitions may be a key mechanism by which CBT exerts its effects, even when direct cognitive interventions are de-emphasized (Lorenzo-Luaces, German, & DeRubeis, 2015).

The skills-based nature of the present protocol suggests the possibility that improved cognitive function might be a critical mechanism of symptom improvement. Impaired cognitive function has been documented in patients with HD via objective neuropsychological testing, albeit with mixed results (see Woody et al., 2014). Here, we focus on subjective cognitive impairment, which may predict functioning and quality of life beyond the

Table 1
Outline of Group Cognitive–Behavioral Therapy for Hoarding Disorder

Session	Session theme	Major group procedures
1	Group introduction	Introduction, group rules, and psychoeducation about hoarding disorder
2	Mechanisms of hoarding	Review of pretreatment photos Psychoeducation about mechanisms of hoarding Goal setting
3	Decision-making and problem-solving	Discussion of scheduling time to sort and discard Discussion about organizational strategies Sorting and discarding practice
4	Decision-making and problem-solving	Guidelines for reducing acquiring Use of virtual store to identify reasons for acquiring Problem-solving training Sorting and discarding practice
5	Emotion regulation	Psychoeducation about emotions Practice identifying emotions with virtual store and items from home
6	Emotion regulation	Psychoeducation about distress tolerance Practice using distress tolerance skills with virtual store and items from home
7	Cognitive restructuring	Psychoeducation about unhelpful thoughts Practice identifying unhelpful thoughts using items from home
8	Cognitive restructuring	Practice using challenging questions for discarding with items from home
9	Motivational enhancement	Psychoeducation about motivation Identification of goals and values Practice sorting and discarding while reviewing personal goals and values
10	Motivational enhancement	Visualization of decluttered spaces Practice sorting and discarding while reviewing personal goals and values
11	Combination	Troubleshooting barriers Sorting and discarding practice
12	Combination	Troubleshooting barriers Sorting and discarding practice
13	Combination	Troubleshooting barriers Sorting and discarding practice
14	Combination	Troubleshooting barriers Sorting and discarding practice
15	Relapse prevention	Review of progress and skills learned Inventorying and challenging negative thoughts about progress Sorting and discarding practice
16	Relapse prevention	Review of posttreatment home photos Skills for maintaining motivation Additional questions End-of-class celebration

contribution of objective neuropsychological test scores (e.g., Hill et al., 2017). HD is associated with a wide range of self-reported cognitive impairments (Samuels et al., 2018), including memory, distractibility, and inattention; these impairments in cognitive function are thought to maintain the avoidance of decision-making characteristic of HD. In particular, distractibility is a promising potential mediator. Impaired capacity to sustain attention is perhaps the most robust neuropsychological finding in HD (Woody et al., 2014), and subjective distractibility shows a larger difference between HD and healthy control participants than do other subjective cognitive functions (Tolin, Hallion, et al., 2018). Distractibility may contribute to difficulty discarding by impeding task persistence; improved focus, conversely, might mediate improved discarding and decreased clutter over the course of treatment. Given the emphasis on improving cognitive efficiency in the present CBT, it would be informative to determine the degree to which improved subjective cognitive function (particularly distractibility, but perhaps other cognitive failures as well) mediates outcomes.

The primary aim of the current study was to determine the efficacy of our redesigned 16-session Group CBT intervention for adults with HD. It was predicted that patients receiving CBT would show a greater reduction on measures of HD severity at posttreatment than would patients assigned to a wait list control condition. The secondary aim was to examine potential mediators of treatment outcome for patients receiving CBT. It was predicted that reductions in both maladaptive cognitions and in subjective cognitive impairment would be significant mediators of HD symptom reduction.

Method

Participants

Eighty-seven adult patients with a primary *DSM–5* diagnosis of HD were enrolled in this study involving the neural mechanisms of CBT response for HD.[1] Prospective participants were primarily recruited from newspaper, radio, and billboard advertisements; online advertising; and community lectures. Participant flow is outlined in Figure 1. Participants were enrolled between October 22, 2013, and

[1] Functional magnetic resonance imaging (fMRI) data for this study will be presented in a separate article.

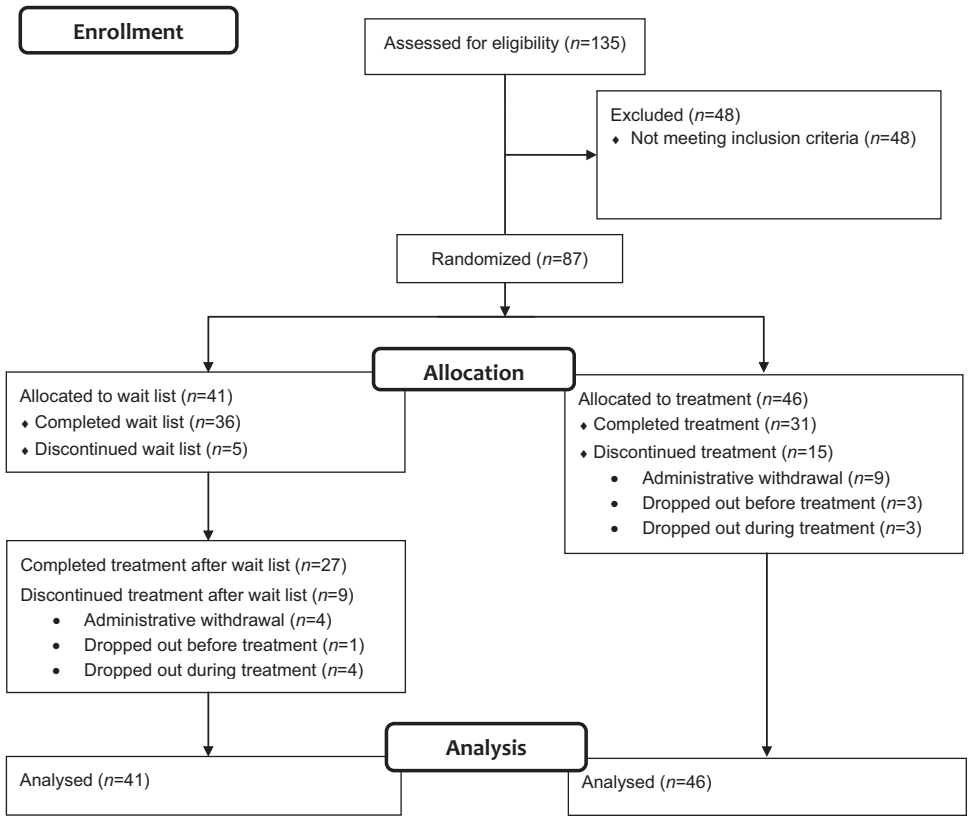

Figure 1. Consolidated Standards of Reporting Trials (CONSORT) diagram.

November 27, 2017. To be included in the study, clinical participants were required to (a) have a primary diagnosis of HD on the *Diagnostic Interview for Anxiety, Mood, and Obsessive-Compulsive and Related Neuropsychiatric Disorders* (Tolin, Gilliam, Wootton, et al., 2018) that was of at least moderate severity according to the *Clinician's Global Impression-HD* (Tolin, Gilliam, Davis, et al., 2018); (b) be aged 20–65; (c) be taking no psychiatric medications or on a stable dose of psychiatric medications (antidepressants, benzodiazepines, and stimulants were allowed) for at least 8 weeks; (d) be willing and able to abstain from the use of stimulant or benzodiazepine medications on the day of testing (though not necessarily during the treatment itself); (e) be free of serious mental illness such as schizophrenia, bipolar disorder, or active substance abuse; (f) be right-handed; and (g) be free of nonremovable metal in the body, claustrophobia, or other factors that would preclude functional magnetic resonance imaging (fMRI). Of the 135 prospective participants that were interviewed, 48 were excluded because of failing to meet inclusion criteria; the most common reasons for exclusion were diagnoses of HD of less than moderate severity ($n = 9$), not a primary diagnosis of HD ($n = 9$), and presence of co-occurring serious mental illness ($n = 9$).

Measures

Primary outcome measure. The primary outcome measure was the *Saving Inventory— Revised* (SI-R; Frost, Steketee, & Grisham, 2004), a 23-item self-report measure that yields a total score as well as three subscales: *Clutter* ($\alpha = .89$ in the present sample), *Saving* ($\alpha = .77$ in the present sample), and *Acquiring* ($\alpha = .88$ in the present sample). The SI-R readily discriminates HD from patients with obsessive–compulsive disorder (OCD) and community controls and correlates significantly with ratings of clutter and impairment (Frost et al., 2004). In the present sample, correlation with the *Hoarding Rating Scale-Interview* (HRS-I) was modest ($r = .50$).

Secondary outcome measures. The HRS-I; Tolin, Frost, & Steketee, 2010) is a 5-item clinician-rated interview of the severity of clutter, difficulty discarding, acquisition, distress, and impairment. The HRS-I has good convergent validity and reliably differentiates HD patients from those with OCD and healthy controls (HCs); (Tolin et al., 2010; Tolin, Gilliam, Davis, et al., 2018). Internal consistency of the HRS-I was borderline in the present sample ($\alpha = .60$), likely because of the small number of items and the fact that each item is intended to measure a dissociable feature of the disorder. Interrater reliability and test–retest reliability in a subset of the present sample (published previously) were good (intraclass correlation coefficients = 0.81 and 0.85, respectively). The *Clinician's Global Impression-Hoarding Disorder* (CGI-HD) scale (Tolin, Gilliam, Davis, et al., 2018) is a variation of the original CGI (Guy, 1976), developed to improve interrater reliability, in which HD symptom severity (based on questions in the HRS-I) is rated on a scale from 1 (*normal, not at all ill*) to 7 (*extremely ill*). A score of 4 (*moderate*) or higher was required for study inclusion. An improvement rating (1–7) was dichotomized into treatment responders (1 = "*very much improved*" or 2 =

"*much improved*") and nonresponders (3 = "*minimally improved*" or higher). The CGI-HD shows good interrater and test–retest reliability (Tolin, Gilliam, Davis, et al., 2018).

Potential mediators. The *Saving Cognitions Inventory* (SCI; Steketee et al., 2003) is a 24-item self-report measure that yields four subscales: emotional attachment (e.g., *"Losing this possession is like losing a friend"*), concerns about memory (e.g., *"Saving this means I don't have to rely on my memory"*), control over possessions (e.g., *"I like to maintain sole control over my things"*), and responsibility toward possessions (e.g., *"I am responsible for the well-being of this possession"*). Respondents rate from 1 to 7 the extent to which they experienced each thought when attempting to discard an object within the past week. The SCI has demonstrated good internal consistency and convergent and discriminant validity in previous samples (Steketee et al., 2003). Internal consistency estimates were acceptable in the present sample (attachment α = .94, memory α = .78, control α = .64, responsibility α = .73). The *Cognitive Failures Questionnaire* (CFQ; Broadbent, Cooper, FitzGerald, & Parkes, 1982) is a 25-item self-report measure assessing the patient's perception of his or her overall cognitive status. Items are rated on a 5-point scale and the CFQ is comprised of four reliable subscales (Wallace, Kass, & Stanny, 2002): Memory (α = .85 in the present sample), Distractibility (α = .83 in the present sample), Blunders (α = .82 in the present sample), and Memory for Names (α = .57 in the present sample).

Other measures. *DSM–5* diagnoses were assessed using the *Diagnostic Interview for Anxiety, Mood, and Obsessive-Compulsive and Related Neuropsychiatric Disorders* (DIAMOND; Tolin, Gilliam, Wootton, et al., 2018), a semistructured clinical interview. The DIAMOND HD diagnosis shows excellent interrater reliability (κ = 0.86), very good test–retest reliability (κ = 0.64), and strong convergence with the Saving Inventory—Revised (Tolin, Gilliam, Wootton, et al., 2018). Primacy of diagnosis was determined using the DIAMOND severity ratings, in which each diagnosis is rated on a scale from 1 (*normal or in remission*) to 7 (*extreme*) based on the participant's level of distress and functional impairment. The *Depression Anxiety Stress Scales* (DASS; Lovibond & Lovibond, 1995) is a 42-item self-report measure assessing three domains of negative emotion: depression, anxiety, and stress/tension. Each item is rated on a 4-point scale assessing symptom frequency over the past week. Subscales of the DASS show high internal consistency (α = .89–0.96) and good discriminant and divergent validity (Lovibond & Lovibond, 1995). In the present study, the depression (DASS-D, α = .94), anxiety (DASS-A, α = .90) and stress (DASS-S, α = .92) subscales showed excellent internal consistency.

Treatment

The CBT followed a published treatment protocol that includes a therapist manual (Tolin, Wootton, Worden, & Gilliam, 2017) and client workbook (Tolin et al., 2017). The group followed a didactic and experiential format for 16 weekly, 90-min sessions. Table 1 outlines the major themes of the treatment and the major content of each session. A class-like group structure was intended to help focus attention on the topic at hand, as well as reduce self-stigma about hoarding behaviors. Two group leaders (a licensed psychologist, who had significant experience working with HD patients, as well as a postdoctoral fellow in clinical psychology) and 4–11 participants per group (M = 7.8) were present for each session. In-session exercises were used for skill acquisition and rehearsal. To practice refraining from acquiring, we created a "virtual store" of 10–20 free items that prior group members had identified as desirable; patients practiced looking at, holding, talking about, and then putting back the items. To practice sorting, decision-making, and discarding, participants were instructed to bring a box or bag of items from home to each session. The group room had boxes labeled "recycle/donate" and "trash" in the middle of the room. Psychoeducation modules involved using the "virtual store" and sorting boxes (e.g., during psychoeducation about emotions, participants were asked to discard an item, or to refrain from acquiring an item, and to identify their emotions). Early in the group, participants were educated about four hypothesized mechanisms of HD: (a) poor decision-making and problem-solving, (b) intense emotions, (c) maladaptive thought processes, and (d) low motivation. Two sessions were dedicated to address each of these mechanisms, with all including psychoeducation and in-session practice of skills. Sessions 11–14 employed all of these themes as patients continued in-session sorting and discarding practice. The final two sessions were dedicated to relapse prevention.

Homework followed each group session and included the completion of assignments and worksheets which helped the patient to review concepts and skills, track possessions going in and out of the home, and monitor between session (at-home) sorting and discarding exercises. Each session began with a homework review and ended with the setting of new discarding goals for the week. Participants were also instructed to take digital pictures of their living spaces at the beginning and end of treatment, which were reviewed in the group on a projector screen.

Several steps were taken to maximize motivation and treatment compliance. First, midway through treatment (where, in our earlier clinical observations, we found motivation often begins to wane), group time was spent using strategies from motivational interviewing (Miller & Rollnick, 2013), including weighing the pros and cons of behavioral change, and reviewing progress on long-term goals and values and comparing these to current behavior. Second, contingency management strategies, as employed in prior HD research (Worden et al., 2017), were used to target treatment attendance and homework completion. For example, participants kept logs of their goals each week (e.g., discard 10 newspapers) and a point was given when each homework assignment was completed and weekly goal met. At 6-week intervals, if 75% of the possible points had been awarded, a group reward was given (e.g., a free movie ticket or group party, but not an object that could add to clutter). The group rewards were earned in all but one group. At an individual level, participants were instructed to brainstorm meaningful, nonacquiring self-rewards (e.g., watch a favorite TV show, go out for ice cream) and contracted to self-administer these rewards on a weekly basis when that week's goal was met.

Finally, we had a written procedure for managing noncompliance with treatment. An episode of noncompliance was defined as a missed session, or not completing more than 30% of the homework assignments for that session, which typically included the In/Out Log, the weekly discarding goal set in collaboration with the group leaders, treatment worksheets, and compliance with instructions to bring in items for in-session discarding decision-making exercises. At the first instance of noncompliance, the group leader reminded the patient of the importance of regular

attendance and/or homework completion. At the second instance of noncompliance, a group leader spoke with the patient before or after session for 5–10 min to reiterate the importance of treatment compliance and to problem-solve any barriers. At the third instance of noncompliance, the patient was required to attend a separate individual meeting (ranging from 15–35 min) with the group leader before being allowed to attend any more group sessions, at which issues in compliance were problem-solved and the patient was inquired as to whether they felt they would be able to comply or whether they would like to discontinue the group at this time. At the fourth instance of noncompliance, the patient was required to meet with the principal investigator before being allowed to attend any more group sessions. If the patient elected to continue at that time, if he or she had a fifth instance of noncompliance, the patient's participation in the group was discontinued.

Procedure

All study procedures were approved by the Hospital Institutional Review Board and all participants provided written informed consent prior to any study procedures. Participants were recruited from the clinic flow, newspaper advertisements, community lectures, and flyers in the community. Participants met with a doctoral-level psychologist or supervised postdoctoral fellow, trained in the use of the DIAMOND, HRS-I, and CGI-HD, to determine whether they met criteria for enrollment. They then completed, on a subsequent day, a battery of self-report measures, including the SI-R, SCI, DASS, and CFQ, along with neuroimaging procedures which will be described elsewhere.

Groups were formed via sequential enrollment and then participants were block randomized (i.e., by group) to either CBT or wait list. The randomization schedule was developed by the study coordinator in consultation with the principal investigator prior to initiating study enrollment using a random number generator. There were no deviations from the randomization schedule. Participants were informed of group assignment by the study research assistant when they were telephoned to schedule the first group meeting. Independent evaluators were not involved in any study procedures related to randomization. The assessors were unaware of participants' treatment condition; awareness was checked by asking raters to guess participants' condition at each visit and to rate their confidence in their guesses. Several procedures were in place to maintain the blind. For example, participants received a verbal reminder at the beginning of each assessment meeting not to disclose information about their treatment condition, group meetings were coded on shared calendars without treatment condition information, and raters were always excused from meetings when information about treatment condition may be discussed. Raters met monthly to corate audio recordings of assessments in order to prevent rater drift. At midtreatment (Week 8) and at posttreatment (Week 16) each participant met with the assessor again, who readministered all study measures. Participants assigned to wait list received no contact for 16 weeks (though they were encouraged to contact the clinic in case of psychiatric emergency), and then received the 16-week CBT protocol. Treatment fidelity was assessed by having an independent psychologist, who was not one of the study clinicians, listen to 10% of audiotaped group sessions. A checklist (available from the authors) was used to determine whether each of the elements of the session manual were present in the session. The trial was registered with clinicaltrials.gov (NCT01956344).

Data Analytic Strategy

Data were sought for all individuals at all time points whether they completed treatment or not for purposes of complete-case (intent-to-treat) analysis. Of the 14 individuals who dropped out of the study during treatment, posttreatment data were obtained for 12 of them. The use of complete case analysis often underestimates treatment efficacy but has the advantage of preserving randomization, which is not necessarily the case for the analysis of just treatment completers, also known as per protocol analysis (Mitka, 2012; National Research Council, 2010). Complete case analyses are reported here to preserve randomization but per protocol analyses that better estimate efficacy presuming dropouts do not disrupt randomization are provided in the online supplemental material. Substantive conclusions were the same for both forms of analysis. Missing data were treated using traditional full information maximum likelihood (FIML) method or, where this was not available, either by chained equation multiple imputation (van Buuren, 2007) or listwise deletion. Confidence intervals (95%) are reported as margins of error representing the absolute value of the half width of the interval relative to the parameter estimate. Cohen's *ds* are reported as absolute values; all were in the predicted direction. The presence of misleading outliers was evaluated in all analyses; none were evident. Most analyses used Huber-White robust estimation to deal with variance heterogeneity and nonnormality and were replicated using percentile bootstrapping (with 1,000 bootstraps), as appropriate, for sensitivity purposes. Conclusions remained the same across both forms of analysis.

Effects on mediator/outcome means as a function of treatment versus wait list control were tested with single degree of freedom contrasts using the corresponding baseline outcome/mediator as a covariate. Given the therapies were administered in small groups, analyses were conducted both with and without cluster adjustments for group membership using algorithms from the Mplus 8.0 computer program (Baldwin, Bauer, Stice, & Rohde, 2011; Muthén & Muthén, 2012). Both cluster unadjusted and adjusted analyses were performed in the spirit of a sensitivity framework because of questions of statistical appropriateness of traditional clustering algorithms given the small number of clusters in the study. None of the reported conclusions differed as a function of these sensitivity analyses, unless otherwise noted.

Because of the small sample size, mediational analyses used limited information structural equation modeling (SEM) rather than full information structural equation modeling (Goldberger, 1972; Jaccard, Guilamo-Ramos, Johansson, & Bouris, 2006). Full information methods estimate linear equations in a path model simultaneously, typically with maximum likelihood algorithms. Limited information estimation works separately with each equation implied by the model, estimating coefficients on a per equation basis but using the overall structural model to guide interpretation and the conduct of supplemental statistical tests. The approach sacrifices traditional global indices of model fit in exchange for flexibility in the use of different estimation methods on a per equation basis and for small sample size appropriateness (see the online supplemental material). For continuous variables in the mediation model, analyses were performed to affirm linearity assumptions for variable

relationships (see online supplemental material). Significance tests of mediation used the joint significance test (MacKinnon, Lockwood, Hoffman, West, & Sheets, 2002), which tends to perform as well as more complex bootstrap methods for mediation analysis (Hayes & Scharkow, 2013).

Finally, to supplement the SEM analysis of mediation, a fixed effects regression analysis for panel data (Allison, 2009) was undertaken using all individuals, including the crossovers from the wait list control after they completed treatment; that is, there was no control group and all individuals had been exposed to treatment. The analysis used three time points, the baseline, a midtreatment assessment (at Week 8), and the posttest. Time was dummy-coded in the analysis. Fixed effects regression analysis removes from the data all sources of between-subjects variability, so any between-subjects confounds (e.g., gender, depression) are moot. The method essentially regresses the outcome variable onto a mediator on a within-person basis, documenting how much the outcome changes given a one unit change in the mediator within an individual, on average.

Results

Sample Description and Participant Flow

As shown in Figure 1, of the 41 patients randomized to wait list, 5 (12%) withdrew from that condition. Of the 46 patients randomized to CBT, 15 (33%) withdrew from that condition: six dropped out on their own accord (three prior to treatment and three during treatment), while nine were removed from treatment because of noncompliance. Of the 36 patients who completed the wait list condition and progressed to CBT, nine subsequently withdrew from that condition: five dropped out on their own accord (one prior to treatment and four during treatment) and four were removed because of noncompliance. Overall, of the 78 patients who started CBT, 7 (9%) dropped out on their own accord and 13 (17%) were removed from the group for noncompliance.

Descriptive statistics for the sample are shown in Table 2. No meaningful pretreatment differences were found on any measure between the CBT and wait list groups, suggesting adequacy of randomization. Scores on the SI-R and HRS-I were well within the clinical range, and the average patient was rated as "markedly ill" on the CGI-HD. Depression scores on the DASS were in the mild range on average, whereas the anxiety and stress scores were not clinically elevated. Scores on all SCI subscales were more than 1 *SD* above the normative means reported in the original SCI publication (Steketee et al., 2003), indicating significant endorsement of maladaptive saving beliefs. The CFQ mean total score was 45.27 ($SD = 15.80$), which is nearly 2 *SD* higher than the healthy control comparison sample used for the fMRI data (results to be reported in a separate paper). Similarly, scores on the Memory, Distractibility, and Blunders subscales were more than 1 *SD* higher in the hoarding patients than the healthy controls. The Memory for Names subscale score was within 1 *SD* of the healthy control group.

Table 2
Baseline Sample Description

Variable	CBT group ($N = 46$)	Wait list group ($N = 41$)
Age, *M* (*SD*)	53.93 (10.12)	54.05 (8.28)
Female, *N* (%)	37 (80.4%)	35 (85.4%)
Nonwhite, *N* (%)	5 (10.9%)	3 (7.3%)
SI-R Total, *M* (*SD*)	61.05 (12.20)	61.87 (9.69)
SI-R Clutter, *M* (*SD*)	25.48 (6.31)	26.31 (4.81)
SI-R Saving, *M* (*SD*)	19.91 (3.68)	19.95 (3.74)
SI-R Acquiring, *M* (*SD*)	15.66 (6.18)	15.62 (4.78)
HRS-I, *M* (*SD*)	27.33 (4.04)	26.93 (4.87)
CGI-HD, *M* (*SD*)	5.07 (.95)	5.11 (.83)
DASS Depression, *M* (*SD*)	8.89 (7.79)	10.64 (9.06)
DASS Anxiety, *M* (*SD*)	5.72 (6.17)	5.92 (7.13)
DASS Stress, *M* (*SD*)	12.95 (8.25)	13.95 (8.71)
SCI Responsibility, *M* (*SD*)	24.49 (8.01)	24.45 (7.11)
SCI Attachment, *M* (*SD*)	41.43 (16.74)	40.33 (15.61)
SCI Memory, *M* (*SD*)	19.59 (7.62)	20.33 (7.27)
SCI Control, *M* (*SD*)	16.00 (3.70)	15.33 (4.06)
CFQ Memory	11.86 (5.94)	11.74 (5.37)
CFQ Distractibility	18.93 (6.55)	18.97 (5.08)
CFQ Blunders	11.45 (5.74)	10.54 (3.70)
CFQ Names	4.84 (1.98)	4.74 (1.74)
Comorbid depressive disorder, *N* (%)	23 (50.0%)	22 (53.7%)
Comorbid anxiety disorder, *N* (%)	17 (37.0%)	14 (34.1%)
Comorbid OCD, *N* (%)	6 (13.0%)	4 (9.8%)
Comorbid ADHD, *N* (%)	8 (17.4%)	7 (17.1%)
On antidepressant medications, *N* (%)	9 (19.6%)	7 (17.1%)
On benzodiazepine medications *N* (%)	5 (10.9%)	2 (4.9%)
On stimulant medications, *N* (%)	3 (6.5%)	3 (7.3%)

Note. CBT = cognitive–behavioral therapy; SI-R = Saving Inventory–Revised; HRS-I = Hoarding Rating Scale-Interview; CGI-HD = Clinician's Global Impression-Hoarding Disorder; DASS = Depression Anxiety Stress Scales; SCI = Saving Cognitions Inventory; CFQ = Cognitive Failures Questionnaire; OCD = obsessive–compulsive disorder; ADHD = attention-deficit/hyperactivity disorder.

Treatment Fidelity and Independent Evaluator Awareness Check

Scores on the fidelity checklist were calculated as a percentage of items checked. Ratings across sessions ranged from 66.7% to 100% ($M = 94.8$, $SD = 11.74$). Thus, fidelity to the treatment protocol appeared excellent. Out of 167 assessments, raters guessed participants' condition correctly 116 (69.5%) times, with 30 (18.0%) instances of reported unblinding by participants (the remainder of guesses were because of symptom change). Mean confidence rating (0–100%) for guesses was 65.27 ($SD = 21.02$).

Efficacy of Group CBT for Hoarding Disorder

Primary outcome. Table 3 presents means and clinically significant change (CSC) percentages for the treatment and control conditions for the SI-R total and subscale scores. We determined CSC by calculating the statistic *C*, representing the cutoff between clinical and normative groups. *C* is calculated as $C = ((M_{HD})(\sigma) + (\mu)(SD_{HD}))/(SD_{HD} + \sigma)$, where M_{HD} and SD_{HD} are the established mean and *SD* for patients with HD, and μ and σ are the normative mean and *SD* (Tolin et al., 2015). We used the HD and normative means provided by Muroff, Underwood, and Steketee (2014), which resulted in the following cutoff values: SI-R total, 43.22; SI-R saving, 14.50; SI-R clutter, 17.89; and SI-R acquiring, 9.92. The differences were statistically significant ($p < .01$) in favor of the CBT (vs. wait list) condition for all major outcomes and all subscales. The standardized mean differences (based on Cohen's *d*) generally were large (1.27 to 1.59) and the percent differences generally were substantial (−4.21 to −16.24). On the

Table 3
Contrasts for Study Outcomes

Outcome	CBT group	Waitlist group	CBT − waitlist difference	Critical ratio	Cohen's *d*
Outcome means at posttreatment					
SI-R Total	42.63	58.87	−16.24 ± 4.78	6.79	1.59 ± .53
SI-R Clutter	18.79	25.44	−6.66 ± 3.68	3.54	1.43 ± .48
SI-R Saving	14.11	19.58	−5.47 ± 1.70	6.28	1.47 ± .52
SI-R Acquiring	9.64	13.84	−4.21 ± 1.50	5.50	1.27 ± .50
HRS-I Total	17.28	25.84	−8.56 ± 2.06	8.27	1.95 ± .67
Outcome percentages at posttreatment					
CGI-HD Responder	83.3	11.1	72.2 ± 18.6	7.75	—
CSC SI-R Total	41.7	8.1	33.6 ± 18.8	3.58	—
CSC SI-R Clutter	41.7	5.4	36.3 ± 18.0	4.02	—
CSC SI-R Saving	52.8	13.5	39.3 ± 20.0	3.91	—
CSC SI-R Acquiring	52.8	27.0	25.7 ± 22.2	2.33	—

Note. CBT = cognitive–behavioral therapy; SI-R = Saving Inventory–Revised; HRS-I = Hoarding Rating Scale-Interview; CGI-HD = Clinician's Global Impression-Hoarding Disorder; CSC = clinically significant change; CBT $n = 46$, Waitlist $n = 41$. Critical ratio (CR) is the parameter estimate divided by its estimated *SE*. CR > 1.96 is $p < .05$ and CR > 2.56 is $p < .01$. Entries after ± are half widths of 95% confidence intervals, analogous to a margin of error.

SI-R, 42% of treated patients met CSC criteria, compared with 8% of waitlist patients. The number needed to treat (NNT) index (Pinson & Gray, 2003) was 2.98.

Secondary outcomes. As shown in Table 3, HRS-I scores significantly favored CBT over wait list, with a large standardized mean difference (1.95). It is noted, however, that only 17% of treated patients (vs. 0% of waitlist patients) met criteria for CSC on the HRS-I. We used the same formula described above for determining CSC, and used the HD and normative means from the original HRS-I validation study (Tolin et al., 2010). This yielded a cutoff score of 13.09 for CSC on the HRS-I. For the CGI-HD scale, 83.3% of the treated individuals were rated as much improved or very much improved, compared with 11.1% in the wait list control group, a rate that is 7.5 times higher.

Mediators of Change in Group CBT

Joint significance test. Table 4 presents group differences in means and percentages for the treatment and control conditions for the proposed mediators as well as their subscales. The group differences for CBT versus control were significant and moderate to large ($d = 0.52$ to 0.86) for all mediators except for the CFQ Names subscale. The CFQ Memory subscale was statistically significant for the unadjusted contrast (critical ratio [CR] = 2.31, $p < .05$) but not for the cluster adjusted contrast (CR = 1.87, $p < .07$). The Cohen's *d* coefficients for statistically significant effects generally were near 0.70.

The primary mediational analyses focused on the total scores of the SCI (maladaptive beliefs about saving) and the CFQ (subjective cognitive impairment). The joint significance test for the relevant structural model (see the online supplemental material) predicts that in addition to significant treatment-control group differences on the mediators reported in Table 4, the path/regression coefficients from SCI and CFQ to the primary outcome variables also should be statistically significant. Huber-White robust regression was performed on each outcome separately using as predictors the SCI, CFQ, the treatment condition (dummy coded), and the baseline measure of the outcome (to increase statistical power in an analysis of covariance sense; see Rausch, Maxwell, & Kelley, 2003). Table 5 presents the relevant regression coefficients, standard error of estimates, and squared multiple

Table 4
Contrasts for Mediators

Mediator means	CBT group	Waitlist group	CBT − waitlist difference	Critical ratio	Cohen's *d*
SCI-Total	79.23	98.32	−19.09 ± 10.70	3.57	.86 ± .48
SCI-Emotion	29.50	37.36	−7.86 ± 5.08	3.03	.71 ± .49
SCI-Memory	16.34	20.65	−4.31 ± 2.55	3.39	.82 ± .49
SCI-Responsibility	19.35	23.77	−4.43 ± 3.06	2.83	.66 ± .49
SCI-Control	13.87	16.14	−2.28 ± 1.69	2.64	.63 ± .50
CFQ-Total	41.26	47.22	−5.96 ± 4.18	2.85	.67 ± .46
CFQ-Memory	10.69	12.55	−1.86 ± 1.60	2.31	.55 ± .50
CFQ-Distraction	16.75	19.61	−2.86 ± 1.81	3.17	.75 ± .46
CFQ-Blunder	10.36	11.85	−1.50 ± 1.30	2.25	.52 ± .45
CFQ-Names	4.56	4.76	−.19 ± .57	.67 (*ns*)	.16 ± .48

Note. CBT = cognitive–behavioral therapy; SCI = Saving Cognitions Inventory; CFQ = *Cognitive Failures Questionnaire*. Critical ratio (CR) is the parameter estimate divided by its estimated *SE*. CR > 1.96 is $p < .05$ and CR > 2.56 is $p < .01$; CBT $n = 46$, Waitlist $n = 41$; entries after ± are half widths of 95% confidence intervals, analogous to a margin of error.

Table 5
Saving Inventory–Revised (SI-R) Based Mediational Regressions

Predictor	SI-R outcome	
	Unstandardized coefficient	Critical ratio
Intercept	16.45 ± 22.04	—
Baseline measure	.58 ± .25	4.57
Treatment group	−14.36 ± 5.78	6.31
SCI	.14 ± .09	2.92
CFQ	−.14 ± .22	1.21
Multiple R^2	.55	—
Standard error of estimate	9.68	—

Note. SCI = Saving Cognitions Inventory; CFQ = *Cognitive Failures Questionnaire*. Critical ratio (CR) is the parameter estimate divided by its estimated standard error. CR > 1.96 is $p < .05$ and CR > 2.56 is $p < .01$. Entries after ± are half widths of 95% confidence intervals, analogous to a margin of error. Total $N = 87$.

correlations. The CFQ did not independently predict SI-R outcomes after the other variables were controlled, which is inconsistent with mediation. The SCI significantly and independently predicted change in SI-R. Treatment condition remained a statistically significant regression coefficient for both hoarding outcomes, suggestive of partial mediation by change in SCI (i.e., SCI accounted for some but not all of the effect of CBT on SI-R). An estimate of how much of the overall treatment effect on SI-R was because of the SCI mediator was 16%, with about 84% of the program effect being because of mechanisms/mediators other than SCI. Tests of treatment by mediator interactions on the outcome using product terms did not yield statistically significant coefficients, suggesting no meaningful treatment by mediator interaction.

Fixed effects panel regression. The fixed effects panel regression analyses tested if changes in the two mediators (entered into the regression equation simultaneously) were associated with changes in each hoarding outcome from baseline to midtreatment to posttreatment on a purely within-individual basis. For the SI-R outcome, there was statistically significant reduction in total SI-R scores between baseline and midtreatment (mean difference = −6.52 ± 2.79, CR = 4.58, $p < .05$) as well as a statistically significant reduction from midtreatment to the immediate posttreatment (mean difference = −6.40 ± 2.65, CR = 4.72, $p < .05$). The coefficient in the model for SCI was 0.30 ± 0.09 (CR = 6.86, $p < .05$), indicating that for every one unit that SCI decreased across time for individuals, SI-R decreased, on average, by 0.30 units, holding constant CFQ. The coefficient for CFQ was 0.09 ± 0.24 (CR = 0.73, *ns*), suggesting that changes in CFQ did not mediate changes in SI-R.

Additional analyses (see online supplemental material) demonstrated that the study was adequately powered, that aggregation bias did not appear to affect results, and that the conclusions of the mediation analysis were reasonable in the presence of modest violations of sequential ignorability (i.e., correlated distrubances between the mediators and outcomes; see Imai, Keele, & Yamamoto, 2010).

Discussion

The first aim of this study was to investigate the efficacy of a 16-session Group CBT intervention for adults with HD in a randomized controlled trial comparing an immediate treatment group to a wait list control condition. It was hypothesized that the Group CBT intervention would outperform the wait list control condition at posttreatment, with large between group effect sizes. The present results provide preliminary support for the efficacy of Group CBT for HD. Between-groups effects were large ($d = 1.59$ for the SI-R Total and 1.95 for the HRS-I). This effect compares favorably to prior controlled trials for HD. For example, Steketee et al. (2010) conducted a trial of individual CBT for HD and after 12 weeks the effect size between CBT and wait list groups was 1.07 on the SI-R and 0.71 on the HRS-I. In another study of a 20-week Group CBT with multiple home visits, Muroff et al. (2012) obtained a difference of $d = 1.03$ on the SI-R but only 0.21 on the HRS-I compared to bibliotherapy. Given the larger effects seen in the present study it is possible that the protocol used results in enhanced outcomes over previously used protocols for both individual and Group CBT for HD. However, this requires further evaluation in future controlled trials.

In the present study 42% of patients met criteria for CSC on the SI-R, which compares favorably to the 35% rate of CSC seen across CBT studies (Tolin et al., 2015), though the majority (58%) of patients were still in the clinical range (lack of CSC) at the conclusion of treatment. Rates of CSC were significantly worse on the HRS-I, with only 17% meeting criteria (and thus 83% still remaining in the clinical range). The HRS-I is likely a more complete (and conservative) measure of outcome, as evidenced by its smaller pre- to posttreatment effects in this and prior research Muroff et al. (2012). This could be because the revised HRS-I used in this study has detailed probe and clarification questions (Tolin, Gilliam, Davis, et al., 2018) that help the interviewer obtain a clearer impression of the severity of symptoms; it is noted that the correlation between the SI-R total score and HRS-I was modest. It is also noted that the HRS-I measures not only the domains tapped by the SI-R (clutter, difficulty discarding, and acquiring) but also hoarding-related distress and impairment; it is possible that these factors attenuated the outcomes as changes in functional impairment frequently lag behind those in the behavioral features of hoarding (Tolin et al., 2015). It is also quite possible that the HRS-I is less sensitive to treatment effects; additional research into this topic is warranted.

There are a number of factors which may enhance the efficacy of CBT for HD. First, as noted previously (Tolin et al., 2015), across studies longer treatment is associated with stronger effects, particularly in the domains of clutter and functional impairment, and it is possible that a longer course of treatment, would be of benefit. Second, enhancing between-session compliance during treatment may also be beneficial given the strong relationship between compliance and outcome (Ayers, Wetherell, Golshan, & Saxena, 2011; Tolin et al., 2007). Given the high rates of attentional deficits (e.g., Grisham, Brown, Savage, Steketee, & Barlow, 2007) and memory confidence concerns (Hartl et al., 2004) seen in this population, automated prompts and reminders may be a pragmatic low-cost way to enhance between-session compliance. While the results from this study provide preliminary support for Group CBT for HD, future research should focus on how to improve this effect further with dismantling studies.

The second aim of the present study was to examine potential mediators of treatment outcome for patients receiving CBT. It was predicted that reductions in both maladaptive cognitions and in

subjective cognitive impairment would be significant mediators of HD symptom reduction. Consistent with prior research using a more cognitively based, manualized, individual CBT (Levy et al., 2017), the findings further suggest a mechanistic role for changes in hoarding-related beliefs in the beneficial effects of CBT on HD, accounting for 16% of treatment effects. This finding is consistent with emerging evidence that cognitive change mediates outcomes even for more behavioral treatments in related disorders (e.g., depression; Lorenzo-Luaces et al., 2015). Thus, it is important for future research to investigate how these hoarding-related beliefs can be more effectively addressed in treatment. The failure of subjective cognitive impairment to mediate treatment outcomes is also notable given the greater emphasis on cognitive skill-building and lesser emphasis on cognitive restructuring, compared with previous CBT protocols. Patients with HD's subjective reports of cognitive functioning do not reliably align with objective measures (Moshier et al., 2016); it therefore remains plausible that objective improvements in cognitive functioning may have contributed to treatment gains despite being undetected by patients. Future research should examine both psychological and biological variables that might further account for symptom improvement. Prior research has begun to link the discarding decision-making process to specific abnormalities in neural activity (e.g., Tolin et al., 2012). Neuroimaging was included in the present study at pre-, mid-, and posttreatment; future analyses should examine whether abnormal patterns of neural activity mediate treatment response as well as hypothesized mediators (e.g., distractibility).

Several important limitations of the present study should be noted. First, the use of a wait list control rather than placebo or *bona fide* treatment precludes strong conclusions about the specificity of the effects to CBT per se (i.e., vs. nonspecific factors). Wait list was deemed to be the most appropriate control condition for the present study given the dearth of randomized controlled trials of CBT for HD in the literature (e.g., only one small controlled trial of Group CBT in which treatment was compared with a bibliotherapy condition; Muroff et al., 2012). However, we note that wait list can serve as a "nocebo" condition, potentially inflating estimates of efficacy (Furukawa et al., 2014). Future studies should compare Group CBT to active control conditions.

Second, there was no follow-up assessment after treatment; therefore, it is unclear whether treatment gains were maintained after treatment discontinuation. This concern is attenuated somewhat by the findings of Muroff, Steketee, Frost, and Tolin (2014), who found generally strong maintenance of treatment gains 1 year after individual CBT for HD. Nevertheless, it is important to examine the long-term efficacy of Group CBT for HD. Such an investigation is particularly important for the current treatment protocol given it differs significantly from treatment delivered in previous trials of CBT for HD, which focused more heavily on traditional cognitive interventions.

Third, though treatment fidelity of prescribed interventions was high there were no systematic ratings made for the inclusion of nonmanual interventions. While the fidelity rater informally noted that these occurred rarely, if at all, in sessions, the lack of a formal quantitative assessment is a limitation of the current study. Fourth, because these data were collected as part of an ongoing clinical trial using fMRI, our inclusion criteria (e.g., medication restrictions, at least moderate HD severity, right-handedness) may not be representative of all treatment-seeking HD patients. The age restriction (65 years or younger, resulting in a mean age of 54 years), used to avoid age-related neuroimaging confounds, is particularly noteworthy: HD is common among older adults (Samuels et al., 2008), and some evidence suggests attenuated treatment response in older HD patients (Ayers et al., 2011). Aging-related decline in executive functioning may introduce unique complications (e.g., exacerbation of impaired decision-making). It will be important to examine whether HD patients over the age of 65 respond to this treatment as well as do those under 65, or whether additional age-related interventions are needed (e.g., Ayers et al., 2014).

Fourth, the unblinding of 18% of IE ratings is another limitation that attenuates the strength of conclusions that can be drawn. Ideally, these raters would have been replaced. In terms of the statistical modeling of mediation, parameter estimates can be biased by endogeneity, specification error, omitted variables, and measurement error, thus requiring caution in conclusions about mediation.

Fifth, we did not administer clutter-specific measures (e.g., the Clutter Image Rating; Frost, Steketee, Tolin, & Renaud, 2008), although we assessed self-reported clutter severity using the clutter subscale of the SI-R. Clutter-specific measures may have provided additional insight into HD symptom change and potential mechanisms of treatment response. We also did not explicitly assess HD-related impairment using an impairment-specific measure. However, we note that impairment is one of the five HRS-I items completed by the IEs.

Sixth, the 26% noncompletion rate is a limitation of the present study. Of the 78 patients who started CBT, 9% dropped out on their own accord and 17% were removed for noncompliance. Other studies of CBT have not included the stepped consequences for noncompliance included in the present protocol, which we thought to be important for maximizing treatment compliance (particularly in a group setting, in which a grossly noncompliant patient could easily affect the compliance of other group members). Therefore, it is problematic to compare the noncontinuation rate with the patient-initiated dropout rates in other studies of HD. Steketee et al.'s (2010) patient-initiated dropout rate of 10% is comparable to the 9% seen here. Mataix-Cols, Marks, Greist, Kobak, and Baer (2002), in a large study of OCD patients, found that 27% of those with significant hoarding symptoms dropped out of behavior therapy, a figure that is quite similar to the noncompletion rate seen here. The pros and cons of removing 17% of patients for noncompliance in the present study is debatable, though we note that this was done only after five instances of noncompliance with repeated warnings.

Finally, as in many clinical trials, the sample was restricted. Our sample was largely White and female, which likely does not reflect the entire HD population in the United States (Samuels et al., 2008). For this reason, it will be important to replicate these findings in more diverse samples of individuals with HD in the future.

The preliminary efficacy of Group CBT demonstrated in the present study suggests a need for both dismantling studies and comparisons with other treatments. Dismantling studies would be valuable as a means to further clarify the active ingredients that are most strongly associated with improvement, which would in turn promote a better understanding of underlying mechanisms. Comparisons with pharmacotherapy, and examinations of the efficacy of combined therapy versus monotherapy, would be useful. The efficacy of pharmacotherapy for HD has not been established in

controlled trials, although some medications, particularly venlafaxine (Saxena & Sumner, 2014) and atomoxetine (Grassi et al., 2016) appear promising in pilot open trials. No non-CBT psychosocial treatments have been tested to date for HD, rendering comparisons to other psychosocial treatment (other than active placebo interventions) premature. Finally, the disseminability of this Group CBT approach to community mental health settings should be examined. Unlike some of the prior work on CBT, the present protocol was designed to be feasible for this application, with no home visits and only 16 weekly sessions. Dissemination efforts have already suggested that many of the principles of CBT can be used successfully by lay group leaders (Frost, Pekareva-Kochergina, & Maxner, 2011; Frost, Ruby, & Shuer, 2012), and a broader effectiveness trial is warranted.

References

Allison, P. (2009). *Fixed effects regression modeling*. Thousand Oaks, CA: Sage. http://dx.doi.org/10.4135/9781412993869

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.

Ayers, C. R., Saxena, S., Espejo, E., Twamley, E. W., Granholm, E., & Wetherell, J. L. (2014). Novel treatment for geriatric hoarding disorder: An open trial of cognitive rehabilitation paired with behavior therapy. *The American Journal of Geriatric Psychiatry, 22,* 248–252. http://dx.doi.org/10.1016/j.jagp.2013.02.010

Ayers, C. R., Wetherell, J. L., Golshan, S., & Saxena, S. (2011). Cognitive-behavioral therapy for geriatric compulsive hoarding. *Behaviour Research and Therapy, 49,* 689–694. http://dx.doi.org/10.1016/j.brat.2011.07.002

Baldwin, S. A., Bauer, D. J., Stice, E., & Rohde, P. (2011). Evaluating models for partially clustered designs. *Psychological Methods, 16,* 149–165. http://dx.doi.org/10.1037/a0023464

Bentler, P. M., & Chou, C. H. (1987). Practical issues in structural modeling. *Sociological Methods & Research, 16,* 78–117. http://dx.doi.org/10.1177/0049124187016001004

Bodryzlova, Y., Audet, J. S., Bergeron, K., & O'Connor, K. (2018). Group cognitive-behavioural therapy for hoarding disorder: Systematic review and meta-analysis. *Health & Social Care in the Community*. Advance online publication. http://dx.doi.org/10.1111/hsc.12598

Bollen, K. (1989). *Structural equations with latent variables*. New York, NY: Wiley. http://dx.doi.org/10.1002/9781118619179

Boomsma, A. (1982). Robustness of LISREL against small sample sizes in factor analysis models. In K. Jöreskog & H. Wold (Eds.), *Systems under indirection observation: Causality, structure, prediction* (Part I; pp. 149–173). Amsterdam, the Netherlands: North-Holland.

Boomsma, A. (1985). Nonconvergence, improper solutions, and starting values in LISREL maximum likelihood estimation. *Psychometrika, 50,* 229–242. http://dx.doi.org/10.1007/BF02294248

Broadbent, D. E., Cooper, P. F., FitzGerald, P., & Parkes, K. R. (1982). The Cognitive Failures Questionnaire (CFQ) and its correlates. *British Journal of Clinical Psychology, 21,* 1–16. http://dx.doi.org/10.1111/j.2044-8260.1982.tb01421.x

Chasson, G. S., Guy, A. A., Bates, S., & Corrigan, P. W. (2018). They aren't like me, they are bad, and they are to blame: A theoretically-informed study of stigma of hoarding disorder and obsessive-compulsive disorder. *Journal of Obsessive-Compulsive and Related Disorders, 16,* 56–65. http://dx.doi.org/10.1016/j.jocrd.2017.12.006

Drury, H., Ajmi, S., Fernández de la Cruz, L., Nordsletten, A. E., & Mataix-Cols, D. (2014). Caregiver burden, family accommodation, health, and well-being in relatives of individuals with hoarding disorder. *Journal of Affective Disorders, 159,* 7–14. http://dx.doi.org/10.1016/j.jad.2014.01.023

Frost, R. O., Pekareva-Kochergina, A., & Maxner, S. (2011). The effectiveness of a biblio-based support group for hoarding disorder. *Behaviour Research and Therapy, 49,* 628–634. http://dx.doi.org/10.1016/j.brat.2011.06.010

Frost, R. O., Rosenfield, E., Steketee, G., & Tolin, D. F. (2013). An examination of excessive acquisition in hoarding disorder. *Journal of Obsessive-Compulsive and Related Disorders, 2,* 338–345. http://dx.doi.org/10.1016/j.jocrd.2013.06.001

Frost, R. O., Ruby, D., & Shuer, L. J. (2012). The Buried in Treasures Workshop: Waitlist control trial of facilitated support groups for hoarding. *Behaviour Research and Therapy, 50,* 661–667. http://dx.doi.org/10.1016/j.brat.2012.08.004

Frost, R. O., Steketee, G., & Grisham, J. (2004). Measurement of compulsive hoarding: Saving inventory-revised. *Behaviour Research and Therapy, 42,* 1163–1182. http://dx.doi.org/10.1016/j.brat.2003.07.006

Frost, R. O., Steketee, G., Tolin, D. F., & Renaud, S. (2008). Development and Validation of the Clutter Image Rating. *Journal of Psychopathology and Behavioral Assessment, 30,* 193–203. http://dx.doi.org/10.1007/s10862-007-9068-7

Furukawa, T. A., Noma, H., Caldwell, D. M., Honyashiki, M., Shinohara, K., Imai, H., . . . Churchill, R. (2014). Waiting list may be a nocebo condition in psychotherapy trials: A contribution from network meta-analysis. *Acta Psychiatrica Scandinavica, 130,* 181–192. http://dx.doi.org/10.1111/acps.12275

Garver, M. S., & Mentzer, J. T. (1999). Logistics research methods: Employing structural equation modeling to test for construct validity. *Journal of Business Logistics, 20,* 33–57.

Gilliam, C. M., Norberg, M. M., Villavicencio, A., Morrison, S., Hannan, S. E., & Tolin, D. F. (2011). Group cognitive–behavioral therapy for hoarding disorder: An open trial. *Behaviour Research and Therapy, 49,* 802–807. http://dx.doi.org/10.1016/j.brat.2011.08.008

Goldberger, A. S. (1972). Structural equation methods in the social sciences. *Econometrica, 40,* 979–1001. http://dx.doi.org/10.2307/1913851

Grassi, G., Micheli, L., Di Cesare Mannelli, L., Compagno, E., Righi, L., Ghelardini, C., & Pallanti, S. (2016). Atomoxetine for hoarding disorder: A pre-clinical and clinical investigation. *Journal of Psychiatric Research, 83,* 240–248. http://dx.doi.org/10.1016/j.jpsychires.2016.09.012

Grisham, J. R., Brown, T. A., Savage, C. R., Steketee, G., & Barlow, D. H. (2007). Neuropsychological impairment associated with compulsive hoarding. *Behaviour Research and Therapy, 45,* 1471–1483. http://dx.doi.org/10.1016/j.brat.2006.12.008

Grisham, J. R., Norberg, M. M., Williams, A. D., Certoma, S. P., & Kadib, R. (2010). Categorization and cognitive deficits in compulsive hoarding. *Behaviour Research and Therapy, 48,* 866–872. http://dx.doi.org/10.1016/j.brat.2010.05.011

Guy, W. (1976). *Assessment manual for psychopharmacology*. Washington, DC: U. S. Government Printing Office.

Hartl, T. L., Frost, R. O., Allen, G. J., Deckersbach, T., Steketee, G., Duffany, S. R., & Savage, C. R. (2004). Actual and perceived memory deficits in individuals with compulsive hoarding. *Depression and Anxiety, 20,* 59–69. http://dx.doi.org/10.1002/da.20010

Hayes, A. F., & Scharkow, M. (2013). The relative trustworthiness of inferential tests of the indirect effect in statistical mediation analysis: Does method really matter? *Psychological Science, 24,* 1918–1927. http://dx.doi.org/10.1177/0956797613480187

Hill, N. L., McDermott, C., Mogle, J., Munoz, E., DePasquale, N., Wion, R., & Whitaker, E. (2017). Subjective cognitive impairment and quality of life: A systematic review. *International Psychogeriatrics, 29,* 1965–1977. http://dx.doi.org/10.1017/S1041610217001636

Hoelter, D. R. (1983). The analysis of covariance structures: Goodness-of-fit indices. *Sociological Methods & Research, 11,* 325–344. http://dx.doi.org/10.1177/0049124183011003003

Imai, K., Keele, L., & Yamamoto, T. (2010). Identification, inference, and sensitivity analysis for causal mediation effects. *Statistical Science, 25,* 51–71. http://dx.doi.org/10.1214/10-STS321

Ivanov, V. Z., Enander, J., Mataix-Cols, D., Serlachius, E., Månsson, K. N. T., Andersson, G., . . . Rück, C. (2018). Enhancing group cognitive–behavioral therapy for hoarding disorder with between-session Internet-based clinician support: A feasibility study. *Journal of Clinical Psychology, 74,* 1092–1105. http://dx.doi.org/10.1002/jclp.22589

Jaccard, J., Guilamo-Ramos, V., Johansson, M., & Bouris, A. (2006). Multiple regression analyses in clinical child and adolescent psychology. *Journal of Clinical Child and Adolescent Psychology, 35,* 456–479. http://dx.doi.org/10.1207/s15374424jccp3503_11

Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59,* 12–19. http://dx.doi.org/10.1037/0022-006X.59.1.12

Kyrios, M. (2014). Psychological models of hoarding. In R. O. Frost & G. Steketee (Eds.), *The Oxford handbook of hoarding and acquiring* (pp. 206–220). New York, NY: Oxford University Press.

Levy, H. C., Worden, B. L., Gilliam, C. M., D'Urso, C., Steketee, G., Frost, R. O., & Tolin, D. F. (2017). Changes in saving cognitions mediate hoarding symptom change in cognitive-behavioral therapy for hoarding disorder. *Journal of Obsessive-Compulsive and Related Disorders, 14,* 112–118. http://dx.doi.org/10.1016/j.jocrd.2017.06.008

Lorenzo-Luaces, L., German, R. E., & DeRubeis, R. J. (2015). It's complicated: The relation between cognitive change procedures, cognitive change, and symptom change in cognitive therapy for depression. *Clinical Psychology Review, 41,* 3–15. http://dx.doi.org/10.1016/j.cpr.2014.12.003

Lovibond, S. H., & Lovibond, P. F. (1995). *Manual for the Depression Anxiety Stress Scales*. Sydney, Australia: The Psychology Foundation of Australia.

MacKinnon, D. P., Lockwood, C. M., Hoffman, J. M., West, S. G., & Sheets, V. (2002). A comparison of methods to test mediation and other intervening variable effects. *Psychological Methods, 7,* 83–104. http://dx.doi.org/10.1037/1082-989X.7.1.83

Mataix-Cols, D., Marks, I. M., Greist, J. H., Kobak, K. A., & Baer, L. (2002). Obsessive-compulsive symptom dimensions as predictors of compliance with and response to behaviour therapy: Results from a controlled trial. *Psychotherapy and Psychosomatics, 71,* 255–262. http://dx.doi.org/10.1159/000064812

Miller, W. R., & Rollnick, S. (2013). *Motivational interviewing: Helping people change* (3rd ed.). New York, NY: Guilford Press.

Mitka, M. (2012). FDA and pharma seek better ways to assess drug safety, efficacy in clinical trials. *Journal of the American Medical Association, 307,* 2576–2577. http://dx.doi.org/10.1001/jama.2012.6684

Moshier, S. J., Wootton, B. M., Bragdon, L. B., Tolin, D. F., Davis, E., DiMauro, J., & Diefenbach, G. J. (2016). The relationship between self-reported and objective neuropsychological impairments in patients with hoarding disorder. *Journal of Obsessive-Compulsive and Related Disorders, 9,* 9–15. http://dx.doi.org/10.1016/j.jocrd.2016.01.002

Mueller, A., Mitchell, J. E., Crosby, R. D., Glaesmer, H., & de Zwaan, M. (2009). The prevalence of compulsive hoarding and its association with compulsive buying in a German population-based sample. *Behaviour Research and Therapy, 47,* 705–709. http://dx.doi.org/10.1016/j.brat.2009.04.005

Muroff, J., Steketee, G., Bratiotis, C., & Ross, A. (2012). Group cognitive and behavioral therapy and bibliotherapy for hoarding: A pilot trial. *Depression and Anxiety, 29,* 597–604. http://dx.doi.org/10.1002/da.21923

Muroff, J., Steketee, G., Frost, R. O., & Tolin, D. F. (2014). Cognitive behavior therapy for hoarding disorder: Follow-up findings and predictors of outcome. *Depression and Anxiety, 31,* 964–971. http://dx.doi.org/10.1002/da.22222

Muroff, J., Steketee, G., Rasmussen, J., Gibson, A., Bratiotis, C., & Sorrentino, C. (2009). Group cognitive and behavioral treatment for compulsive hoarding: A preliminary trial. *Depression and Anxiety, 26,* 634–640. http://dx.doi.org/10.1002/da.20591

Muroff, J., Underwood, P., & Steketee, G. (2014). *Group treatment for hoarding disorder: Therapist Guide*. New York, NY: Oxford University Press. http://dx.doi.org/10.1093/med:psych/9780199340965.001.0001

Muthén, L. K., & Muthén, B. O. (2012). *Mplus User's Guide* (7th ed.). Los Angeles, CA: Author.

National Research Council. (2010). *Panel on handling missing data in clinical trials: The prevention and treatment of missing data in clinical trials*. Washington, DC: National Academies Press.

Nordsletten, A. E., Reichenberg, A., Hatch, S. L., Fernández de la Cruz, L., Pertusa, A., Hotopf, M., & Mataix-Cols, D. (2013). Epidemiology of hoarding disorder. *The British Journal of Psychiatry, 203,* 445–452. http://dx.doi.org/10.1192/bjp.bp.113.130195

Pinson, L., & Gray, G. E. (2003). Psychopharmacology: Number needed to treat: An underused measure of treatment effect. *Psychiatric Services, 54,* 145–146, 154. http://dx.doi.org/10.1176/appi.ps.54.2.145

Rausch, J. R., Maxwell, S. E., & Kelley, K. (2003). Analytic methods for questions pertaining to a randomized pretest, posttest, follow-up design. *Journal of Clinical Child and Adolescent Psychology, 32,* 467–486. http://dx.doi.org/10.1207/S15374424JCCP3203_15

Samuels, J. F., Bienvenu, O. J., Grados, M. A., Cullen, B., Riddle, M. A., Liang, K. Y., . . . Nestadt, G. (2008). Prevalence and correlates of hoarding behavior in a community-based sample. *Behaviour Research and Therapy, 46,* 836–844. http://dx.doi.org/10.1016/j.brat.2008.04.004

Samuels, J., Bienvenu, O. J., Krasnow, J., Wang, Y., Grados, M. A., Cullen, B., . . . Nestadt, G. (2018). Self-reported executive function and hoarding in adults with obsessive-compulsive disorder. *Comprehensive Psychiatry, 81,* 53–59. http://dx.doi.org/10.1016/j.comppsych.2017.11.009

Saxena, S., & Sumner, J. (2014). Venlafaxine extended-release treatment of hoarding disorder. *International Clinical Psychopharmacology, 29,* 266–273. http://dx.doi.org/10.1097/YIC.0000000000000036

Shaw, A. M., & Timpano, K. R. (2016). An experimental investigation of the effect of stress on saving and acquiring behavioral tendencies: The role of distress tolerance and negative urgency. *Behavior Therapy, 47,* 116–129. http://dx.doi.org/10.1016/j.beth.2015.10.003

Steketee, G., & Frost, R. O. (2014). *Compulsive hoarding and acquiring: Therapist guide* (2nd ed.). New York, NY: Oxford University Press.

Steketee, G., Frost, R. O., & Kyrios, M. (2003). Cognitive aspects of compulsive hoarding. *Cognitive Therapy and Research, 27,* 463–479. http://dx.doi.org/10.1023/A:1025428631552

Steketee, G., Frost, R. O., Tolin, D. F., Rasmussen, J., & Brown, T. A. (2010). Waitlist-controlled trial of cognitive behavior therapy for hoarding disorder. *Depression and Anxiety, 27,* 476–484. http://dx.doi.org/10.1002/da.20673

Tolin, D. F., Frost, R. O., & Steketee, G. (2007). An open trial of cognitive-behavioral therapy for compulsive hoarding. *Behaviour Research and Therapy, 45,* 1461–1470. http://dx.doi.org/10.1016/j.brat.2007.01.001

Tolin, D. F., Frost, R. O., & Steketee, G. (2010). A brief interview for assessing compulsive hoarding: The Hoarding Rating Scale-Interview. *Psychiatry Research, 178,* 147–152. http://dx.doi.org/10.1016/j.psychres.2009.05.001

Tolin, D. F., Frost, R. O., Steketee, G., Gray, K. D., & Fitch, K. E. (2008). The economic and social burden of compulsive hoarding. *Psychiatry Research, 160,* 200–211. http://dx.doi.org/10.1016/j.psychres.2007.08.008

Tolin, D. F., Frost, R. O., Steketee, G., & Muroff, J. (2015). Cognitive behavioral therapy for hoarding disorder: A meta-analysis. *Depression and Anxiety, 32,* 158–166. http://dx.doi.org/10.1002/da.22327

Tolin, D. F., Gilliam, C. M., Davis, E., Springer, K., Levy, H. C., Frost, R. O., . . . Stevens, M. C. (2018). Psychometric properties of the Hoarding Rating Scale-Interview. *Journal of Obsessive-Compulsive and Related Disorders, 16,* 76–80. http://dx.doi.org/10.1016/j.jocrd.2018.01.003

Tolin, D. F., Gilliam, C., Wootton, B. M., Bowe, W., Bragdon, L. B., Davis, E., . . . Hallion, L. S. (2018). Psychometric properties of a structured diagnostic interview for *DSM–5* anxiety, mood, and

obsessive-compulsive and related disorders. *Assessment, 25,* 3–13. http://dx.doi.org/10.1177/1073191116638410

Tolin, D. F., Hallion, L. S., Wootton, B. M., Levy, H. C., Billingsley, A. L., Das, A., . . . Stevens, M. C. (2018). Subjective cognitive function in hoarding disorder. *Psychiatry Research, 265,* 215–220. http://dx.doi.org/10.1016/j.psychres.2018.05.003

Tolin, D. F., Stevens, M. C., Villavicencio, A. L., Norberg, M. M., Calhoun, V. D., Frost, R. O., . . . Pearlson, G. D. (2012). Neural mechanisms of decision making in hoarding disorder. *Archives of General Psychiatry, 69,* 832–841. http://dx.doi.org/10.1001/archgenpsychiatry.2011.1980

Tolin, D. F., Wootton, B. M., Worden, B. L., & Gilliam, C. M. (2017). *CBT for hoarding disorder: A group therapy program: Client workbook*. West Sussex, United Kingdom: Wiley-Blackwell. http://dx.doi.org/10.1002/9781119159278

Tryon, W. W. (2018). Mediators and mechanisms. *Clinical Psychological Science, 6,* 619–628.

van Buuren, S. (2007). Multiple imputation of discrete and continuous data by fully conditional specification. *Statistical Methods in Medical Research, 16,* 219–242. http://dx.doi.org/10.1177/0962280206074463

Wallace, J. C., Kass, S. J., & Stanny, C. J. (2002). The cognitive failures questionnaire revisited: Dimensions and correlates. *The Journal of General Psychology, 129,* 238–256. http://dx.doi.org/10.1080/00221300209602098

Woody, S. R., Kellman-McFarlane, K., & Welsted, A. (2014). Review of cognitive performance in hoarding disorder. *Clinical Psychology Review, 34,* 324–336. http://dx.doi.org/10.1016/j.cpr.2014.04.002

Woolridge, J. (2019). *Introductory econometrics: A modern approach* (7th ed.). Boston, MA: Cengage Learning.

Worden, B. L., Bowe, W. M., & Tolin, D. F. (2017). An open trial of cognitive behavioral therapy with contingency management for hoarding disorder. *Journal of Obsessive-Compulsive and Related Disorders, 12,* 78–86. http://dx.doi.org/10.1016/j.jocrd.2016.12.005

Received August 20, 2018
Revision received February 4, 2019
Accepted March 7, 2019 ■