A. Title: Comparing Peer-Led Support Groups with Therapist-Led Support Groups for Treating Hoarding Disorder

Kevin L. Delucchi, PhD¹, Carol A Mathews, MD², R. Scott Mackin, PhD¹, Chia-Ying Chou, PhD¹, Soo Y Uh, PhD¹, L. David Bain, MS³, Sandra J Stark³, Michael Gause, MFA⁴, Ofilio R Vigil, MS⁶, John Franklin, MA³, Mark Salazar, BA³, Julian Plumadore³, Lauren C Smith, MS¹, Kiya Komaiko, BA², Gillian Howell, BA³, Eduardo Vega, MA⁵, Joanne Chan, PsyD³, Monika B Eckfield, PhD⁷, Janice Y Tsoh, PhD¹

¹University of California San Francisco
²University of Florida
³Mental Health Association San Francisco
⁴Sonoma County Community Development Commission
⁵Dignity Recovery Action International
⁶University of California Davis
⁷California State University Eastbay

Original Project Title: Comparison of Peer-Facilitated Support Group and Cognitive Behavioral Therapy for Hoarding Disorder

PCORI ID: CE-1304-6000
HSRProj ID: 20143536
ClinicalTrials.gov ID: NCT02297815

Table of Contents

Abstract ............................................................................................................................................. 3
Background ......................................................................................................................................... 5
Participation of Patients and Other Stakeholders in the Design and Conduct of Research and
Dissemination of Findings .................................................................................................................. 9
  Types and Number of Stakeholders Involved .................................................................................... 9
  How the Balance of Stakeholder Perspectives Was Conceived and Achieved .................................... 10
  Methods Used to Identify and Recruit Stakeholder Partners ........................................................... 10
  Methods, Modes, and Intensity of Engagement ............................................................................... 11
  Perceived or Measured Impact of Engagement .............................................................................. 11
Methods .......................................................................................................................................... 13
  Study Design .................................................................................................................................... 13
  Forming the Study Cohort ................................................................................................................ 13
  Study Setting .................................................................................................................................... 17
  Interventions ..................................................................................................................................... 19
  Follow-up ......................................................................................................................................... 20
  Study Outcomes .............................................................................................................................. 21
  Data Collection and Sources ............................................................................................................ 21
  Analytical and Statistical Approaches ............................................................................................. 25
  Conduct of the Study ........................................................................................................................ 28
Results ............................................................................................................................................. 29
  Baseline (Pretreatment) Characteristics ............................................................................................ 32
  Participation in Treatment ................................................................................................................ 34
  Primary Treatment Outcomes .......................................................................................................... 35
  Secondary Treatment Outcomes ....................................................................................................... 38
Discussion ....................................................................................................................................... 46
  Decisional Context ............................................................................................................................ 46
  The Study Results in Context .......................................................................................................... 46
  Implementation of Study Results ..................................................................................................... 48
  Generalizability ............................................................................................................................... 49
  Subpopulation Considerations ......................................................................................................... 49
  Study Limitations ............................................................................................................................ 50
  Future Research .............................................................................................................................. 51
Conclusion ....................................................................................................................................... 52
References ....................................................................................................................................... 53
Publications ...................................................................................................................................... 58
Appendix ......................................................................................................................................... 59
Abstract

Background: Hoarding disorder (HD) has a profound public health burden and affects about 2% to 6% of the population, but treatment options are limited. The standard of care is cognitive behavioral therapy (CBT), but access to CBT is limited. Individuals may not seek treatment for multiple reasons, including lack of trained providers, poor insight into their illness, stigma, unwillingness to seek help from mental health providers, or cost restrictions. Alternative forms of treatment, particularly community-based treatments, are needed. Preliminary evidence suggests that manual-based peer-facilitated treatment (PFT) for HD may be as effective as CBT; however, PFT has never been directly compared with CBT.


Methods: Three hundred and twenty-three adults with HD were randomly assigned to CBT or PFT. HD severity was assessed at baseline, immediately after treatment (6 months postbaseline), and ≥ 3 months posttreatment, then examined using a test of noninferiority (1-tailed t test). Participants received manualized group CBT led by psychologists (N = 160) or workbook-based group PFT led by peer facilitators (N = 163). The primary outcome was posttreatment hoarding symptom severity (Saving Inventory-Revised [SI-R]; score range 0-92). Secondary outcomes included hoarding-related functional impairment, symptom severity at ≥ 3 months, response (SI-R change ≥ 10 points), and remission (SI-R change ≥ 14 points and SI-R score < 42 points at the posttreatment assessment).

Results: Of participants, 74.6% were female and 59% were white. The mean age was 59 (SD = 10.7). Of randomized participants, 71.5% (N = 231) completed treatment, 247 provided follow-up data (including 16 noncompleters), and 183 (56.7%) provided longitudinal data (mean 14 months posttreatment; range 3-25 months). Dropout rates did not differ between CBT and PFT. CBT participants attended more sessions (mean = 73.3% of groups; SD = 34.0) than did PFT participants (mean = 57.7% of groups; SD = 34.6) (t = −4.09; df = 1; p < 0.0001). The test of the
hypothesis assuming inferiority of PFT to CBT at posttreatment on SI-R scores was rejected (mean difference between conditions = 1.82 points; \( t = -1.71; p = 0.04 \)), indicating no significant difference between groups. Mean pretreatment and posttreatment SI-R scores for the entire sample were 65.6 (SD = 11.7) and 46.8 (SD = 14.6), respectively, equaling a 17.8-point (26.7%) symptom reduction (\( t = 18.96; df = 1; p < 0.00001 \)). Of participants, 66% responded to treatment (SI-R reduction ≥ 10 points). Treatment gains were maintained at 3 months posttreatment. Most participants expressed no strong preference for either treatment, and preference did not affect treatment outcome. More severe hoarding severity at baseline was the only participant-related variable that independently predicted improved treatment outcome for CBT. For PFT, higher pretreatment hoarding severity and better homework adherence were significant predictors of improvement. Multiple measures of neurocognitive function improved after treatment.

**Conclusions:** Group PFT is as effective as group CBT for the treatment of HD, providing an alternative avenue of care for individuals with this functionally impairing disorder.

**Limitations and subpopulation considerations:** Longitudinal follow-up assessment was not planned prior to study initiation but was developed after data collection had started. Subpopulations were not formally examined, although there were no differences in treatment outcome by sex, age, race, insurance status, or psychiatric comorbidity
Background

Hoarding behaviors, which have been described in the scientific literature for more than 35 years,\(^1\) are widespread in the general population, and may be adaptive at low levels or under certain circumstances.\(^2,3\) Hoarding behaviors frequently occur in the context of other neuropsychiatric disorders, including obsessive-compulsive disorder (OCD), schizophrenia, autism, and dementia,\(^4,5\) but pathological or problematic hoarding also occurs independently of these disorders,\(^6\) and was first described in the early 1990s,\(^7\) but not formally recognized by the psychiatric community as an independent psychiatric disorder until 2013.\(^8\)

Hoarding disorder (HD) is defined in the *Diagnostic and Statistical Manual of Mental Disorders*, 5th edition (DSM-5) as persistent difficulty in discarding or parting with items, regardless of their actual value, due to a perceived need to save items and to distress associated with discarding them.\(^8\) Difficulty discarding leads to functional (including social and occupational) impairment because the accumulation of clutter makes active living or work areas unusable for their normal purposes, and in severe cases can lead to difficulty in maintaining a safe living environment. Some, but not all, individuals also have excessive acquiring behaviors, which contributes to the accumulation of clutter. Insight ranges from good, where the individual recognizes that hoarding beliefs or behaviors are problematic, to absent, with delusional beliefs, where the individual is completely convinced that hoarding-related beliefs and behaviors are not problematic despite evidence to the contrary.\(^8\)

Although underrecognized and underdiagnosed, problematic hoarding is quite common, and has a profound impact on functioning and quality of life for affected individuals and their families.\(^9,10\) The only study to assess the epidemiology of DSM-5–defined HD found an estimated prevalence of 1.5%.\(^11\) Two additional population-based studies that used self-report measures to approximate the prevalence of HD using DSM-5 criteria found prevalence rates of 2%\(^12\) and 3.7% (with a weighted population estimate of 5.3%).\(^13\) Some studies suggest that prevalence of problematic hoarding increases substantially among individuals older than age 55, with rates as high as 6%.\(^12-14\)
HD is a chronic disorder, and is associated with high levels of distress, social disruption, and maladjustment (eg, low marriage rates, high social withdrawal), with levels of depression and functional impairment that exceed those of OCD and anxiety disorders. Cluttered homes due to hoarding lead to safety hazards; increased physical morbidity and mortality; difficulty with self-care; and social, financial, and familial consequences. Each year, public service agencies expend tremendous time and financial resources on HD; in San Francisco alone in 2009, more than 6 million dollars per year was spent by service agencies and landlords on hoarding-related issues (not including treatment).

In addition to the burden of multiple chronic cooccurring neuropsychiatric disorders such as anxiety and depression, individuals with HD may also have neurocognitive changes or problems. Neuropsychological studies suggest that executive dysfunction is a prominent feature of HD, characterized primarily by dysfunction in visual learning and categorization, problem solving, information processing speed, and visual memory. Although the association has not yet been directly examined, it is possible that poor adherence to HD-specific treatments or poor outcomes of such treatments is associated with neurocognitive dysfunction, as is seen in other neuropsychiatric disorders such as late-life depression, and that individuals with specific forms of neurocognitive dysfunction may have more difficulties accessing or using specific treatment mechanisms.

Few treatments are currently available for HD, and only a small number of research studies have examined treatment outcome. Four systematic reviews or meta-analyses have been published, summarizing the currently available data on treatment of pathological hoarding and HD. These studies suggest that HD is similar to many chronic neuropsychiatric disorders in that the result of treatment is often improvement of symptoms rather than remission. Although there is a role for pharmacological treatment of HD, at the present time, cognitive behavioral therapy (CBT), typically delivered by trained mental health professionals (eg, psychologists, social workers, psychiatrists) has the strongest evidence base. CBT can be delivered individually or in a group setting, and both settings have been examined for efficacy in HD, usually compared with a waitlist control. 

In part because of the intensive and specialized nature of CBT for HD, and the limited
number of trained treatment providers, a manualized self-help approach was developed, using a workbook called *Buried in Treasures* (BiT),\(^5\) designed for use either by individuals or with facilitators in a group context (peer-facilitated treatment, or PFT).\(^3\) Peer-facilitated interventions have a long history, in particular with individuals with severe mental illnesses such as schizophrenia and bipolar disorder, and as described in a recent systematic review typically fall into 3 types: peers added to traditional mental health services (most common), peers providing clinical services in lieu of more traditional mental health providers (least common), and peers delivering structured curricula (becoming more common).\(^5\) As outlined in a review by Chinman et al, the effectiveness of peer support systems varies widely, with the strongest data for a beneficial outcome coming from the peers added to existing services and the most mixed data arising from studies investigating peers providing clinical services in an established role.\(^5\) Although promising, few studies have been conducted on the efficacy of PFT-based approaches for hoarding.\(^3,3\) In general, however, meta-analyses suggest that individual CBT, group CBT, and group PFT have similar efficacies in HD, while individual self-help approaches, including individual use of the BiT workbook, appear to be less effective.\(^4\) Unfortunately, the sample sizes in these studies are quite small, and the studies are limited in number, leading to variability in outcomes (eg, average pretreatment to posttreatment change scores, percent showing clinically significant improvement) (see Mathews et al for a review).\(^4\) Thus, the degree of improvement that can be expected from CBT or PFT has been difficult to assess, and attempts to examine specific predictors of response, such as clinical, demographic, neurocognitive, or other individual characteristics, have been of limited utility.

HD is underrecognized within the health care community\(^57\) and, outside of a few specialized centers, few treatment options exist. Although HD-focused CBT can be considered the present standard of care treatment for HD, few health care providers are trained to offer it, and there is a pressing need for additional treatment options. Additional barriers to treatment may include lack of insight into illness, stigma, unwillingness to seek help in a mental health setting, or the cost of HD-specific mental health treatment. Psychotherapeutic or other community-based approaches (including peer-facilitated support groups) may be a more inviting alternative for some individuals and their families. To date, there has been only one
direct comparison of the 2 most promising and most easily disseminated forms of treatment, group CBT and group PFT, and the study conducted by our group was hampered by methodological problems. This relative lack of data makes it difficult for individuals, families, health care providers, or health care systems to assess the relative efficacy and fit of these hoarding-specific treatments. The current study aimed to remedy that lack by conducting a clinical trial comparing group CBT against group PFT, which can be much more widely implemented in the general community than CBT, as it does not require trained health care providers. We chose group over individual treatment because it allows for the treatment of a larger number of individuals in a given time period. In addition, we aimed to determine which clinical, demographic, or other individual characteristics (eg, sex, cooccurring depression or anxiety, specific neurocognitive dysfunction) are most predictive of treatment adherence and response for each treatment modality (ie, CBT or PFT).

The aims of the study were as follows:

Aim 1 (primary): Compare the effectiveness of manual-based group CBT and PFT for the treatment of HD in a randomized clinical trial.

Hypothesis: Group PFT will be as effective in treating HD as group CBT.

Aim 2 (secondary): Identify demographic and clinical characteristics predicting treatment response and adherence.

Hypothesis: Clinical and neuropsychological characteristics, including psychiatric comorbidity and impairments in categorization and visual learning and memory, will be associated with a worse outcome. Individuals with higher rates of neuropsychological or psychiatric impairment will be more likely to respond to CBT than to PFT.

Aim 3 (secondary): Determine whether strong preferences for a specific treatment type are associated with treatment adherence and outcome.

Hypothesis: Participants matched to their preferred treatment will do better than those who are not matched to their preferred treatment.


Hypothesis: Measures of visual learning and memory, categorization, and information processing will improve with treatment.
Participation of Patients and Other Stakeholders in the Design and Conduct of Research and Dissemination of Findings

Types and Number of Stakeholders Involved

Relevant stakeholders were engaged at all phases of the research, including planning and design of the study, implementation of study protocol procedures, and monitoring and evaluating treatment outcomes. The stakeholders with the most engagement in these activities were the individuals on the research team and the governance board (described below). Additional stakeholders were invited to give input in a yearly fashion through feedback during presentation of the project and results at the International Conference on Hoarding and Cluttering held by the Mental Health Association of San Francisco (MHASF). The stakeholders in both components were demographically diverse and included individuals from a variety of age groups (ranging from 20s to 80s), racial and ethnic backgrounds, socioeconomic backgrounds, sexual orientations, and relationships to hoarding disorder (eg, family members, caregivers, case managers, public service agencies, peer facilitators, clinicians and mental health professionals, community treatment providers, patient advocate representatives).

Researchers (Drs. Carol Mathews, Scott Mackin, and Kevin Delucchi), consumer advocates with a specific interest in HD (Mr. Eduardo Vega and Mr. Michael Gause), community and academic health care providers who treat HD (Drs. Joanne Chan, Monika Eckfield, and Mathews), peer facilitators (Mr. John Franklin and Mr. Julian Plumadore), and individuals with lived experience of hoarding challenges (Mr. Plumadore) designed this study in a collaborative fashion. Individuals with lived experience of hoarding (Mr. David Bain and Ms. Sandra Stark) were recruited by MHASF as the peer facilitators. The governance board consisted of the entire research team (Dr. Mathews, Dr. Delucchi, Dr. Mackin, Dr. Soo Uhm, Dr. Chia-Ying Chou, Mr. Ofilio Vigil, Ms. Lauren Smith, Ms. Kiya Komaiko from the University of California, San Francisco (UCSF), Mr. Gause, Mr. Vega, Mr. Mark Salazar, Ms. Gillian Howell, Ms. Stark, Mr. Bain, Dr. Eckfield, Dr. Chan, Mr. Plumadore, and Mr. Franklin from MHASF), plus 2 additional individuals, one with lived experience of hoarding (Dr. Fred Lipschultz) and a representative (Ms. Jill Nielson) from Adult Protective Services, which has a specific interest in older individuals with HD. The research team met monthly to discuss the progress of the study and to problem solve as questions or concerns arose. The additional members of the governance board attended at
least quarterly. The role of the additional governance board members was to provide perspective on progress and process of the conduct of the research from outside of the immediate research team and to help solve any dilemmas that could not be addressed by the research team (eg, potential lags in recruitment of participants).

**How the Balance of Stakeholder Perspectives Was Conceived and Achieved**

Stakeholder perspectives were conceived and achieved mainly through the input and assistance from our partnership with MHASF, a nonprofit consumer-run organization that provides leadership and advances the mental health of the people of San Francisco, including those with HD, and leads the global community in advocacy, education, and research in the areas of HD and mental health stigma more generally. MHASF brings diverse stakeholders together and fosters collaboration between disparate viewpoints, as well as brings people and families challenged by mental illnesses together with the agencies that serve them to promote prevention, access to services, leadership, recovery, cultural competence, and independence. MHASF was an equal partner in all aspects of the research.

Although the MHASF team did not have a separate research project within this application, the components of the proposal (and the budget for the work) were divided between the 2 participating sites according to their specific strengths. The MHASF team contained administrative staff, peer leaders, and research staff with expertise in hoarding-related programs, including individuals who oversee the annual MHASF Conference on Hoarding and Cluttering. The MHASF team members were particularly suited for their role in this project, which would not have been possible without their collaboration. The UCSF team consisted of a psychiatrist, a biostatistician, a neuropsychologist with expertise in neurocognitive assessment of individuals with HD, research coordinators, and postdoctoral psychology fellows trained to deliver CBT following the HD-specific CBT manual. The collaborative relationship between the investigators at UCSF and those at MHASF complemented each other’s unique specialties and strengths.

**Methods Used to Identify and Recruit Stakeholder Partners**

We were able to identify and recruit stakeholder partners due largely in part to MHASF, as it specializes in convening stakeholders regarding critical policy and advocacy issues related
to mental illness. In 2007, Dr. Mathews was invited to be a part of MHASF’s San Francisco Task Force on Compulsive Cluttering, which led to the development of the relationship between MHASF and UCSF. Additionally, MHASF’s Annual Conference on Hoarding and Cluttering (now an international event entering its 18th year) brings individuals from around the world together for learning and collaboration and averages close to 400 participants annually.

**Methods, Modes, and Intensity of Engagement**

The stakeholders described above were engaged throughout the entire project period, each group taking on separate responsibilities based on expertise. UCSF took primary responsibility for clinical and neuropsychological assessments, data management, and analyses. The UCSF team was also responsible for the day-to-day project management, including ensuring that targets and milestones were met, preparing reports and other relevant documentation, and overseeing and assessing any adverse events or clinical emergencies. The MHASF team took primary responsibility for participant recruitment, training the CBT and PFT group leaders, scheduling the group interventions, supervising the group leaders, and running the annual hoarding-related conference. UCSF participated in the annual conference, which provides information and education to clients, families, and health care professionals, and continues to be a key source of results dissemination to multiple groups of stakeholders in the San Francisco Bay Area. All members of the research team met monthly at the MHASF offices, joined by the 2 additional outside governance board members on a quarterly basis. The purpose of these meetings was to review the progress of the work, discuss and approve any needed changes to the protocol, and review updates and other documents for the Data and Safety Monitoring Board. Study results, both interim analyses and final analyses, were disseminated to the research team and outside governance board members prior to publication/submission.

**Perceived or Measured Impact of Engagement**

Through collaboration with relevant stakeholders and individuals with lived experience, we have ensured that the research questions, study design, and dissemination of findings are appropriate and accessible to individuals with HD, and that the proposed outcomes are relevant and valuable to those affected by HD. For example, because hoarding symptom
severity, measured by the Saving Inventory-Revised (SI-R), is the metric most frequently used to assess symptom improvement but may not be the outcome of most importance to individuals with HD, we also measured improvement in day-to-day functioning and quality of life at the suggestion of the MHASF team. Similarly, change in thoughts (cognitions) about hoarding behaviors may be the first indicator of later symptom improvement, indicating improved ability to make use of CBT. Therefore, we measured cognitions about hoarding behaviors pretreatment and posttreatment, as such information may be of relevance to health care providers, particularly those administering CBT for HD.

Further, since people with lived experience of hoarding challenges were engaged in these processes, we were able to apply more sensitive and appropriate terminology (ie, “people with lived experience” versus “people with hoarding disorder”) in our recruitment, potentially making future participants feel more comfortable inquiring into the study. We also made additional changes to aspects of the treatments at the request of the MHASF team. For example, while the CBT treatment groups, which were led by trained clinicians, had standardized predefined group treatment agreements and home visits as a part of their curriculum, the peer-led PFT treatment leaders chose not to perform home visits in their curriculum due to the potential discomfort for them and the participants since they are not trained clinicians. Additionally, the PFT treatment leaders did not use the same group agreement documents as the CBT leaders, but created a set of rules/agreements—a “comfort agreement”—developed with each PFT group individually. These comfort agreements generally paralleled the group agreement documents used by the CBT groups, but the rules were decided by consensus of the group members, rather than being predetermined by the leaders. The PFT leaders felt this was a more appropriate method given their relationship with hoarding. We specifically designed the setting of this study to replicate a real-world environment, in that although similar in length and number of sessions, the CBT and PFT groups were not completely parallel to each other but were conducted how they would be regardless of the research component.
Methods

Study Design

We designed the study as a stratified, randomized, single-blind trial. Individuals were stratified by sex and randomized to participate in either CBT or PFT. The study biostatistician (Dr. Delucchi) randomized individuals in a 1:1 ratio in blocks of 5 using a computerized randomization method and provided this information to the UCSF research coordinator, who assigned the treatments after enrollment. Participants were not blinded to group assignment. Study personnel conducting the clinical and neuropsychological assessments at baseline, posttreatment, and at the longitudinal follow-up were blinded to treatment group randomization. Group leaders were blinded to clinical and neurocognitive status of the participants at baseline in order to prevent the introduction of unconscious bias in delivering treatment. We designed the study as a noninferiority trial, testing the hypothesis that PFT would be as effective than CBT, as our preliminary data\textsuperscript{49} and outcomes previously reported for CBT and PFT\textsuperscript{48} suggested that these treatments have similar efficacies.

Forming the Study Cohort

The study was registered with clinicaltrials.gov (NCT02040805) prior to starting any recruitment. Individuals with HD were recruited through (1) advertisements on radio and the internet (eg, the MHASF and International Obsessive Compulsive Disorder Foundation [IOCDF] websites, hoarders.org); (2) listservs of the MHASF and the Bay Area chapter of the IOCDF; (3) the MHASF staff calling or emailing social service agencies, professional organizers, and therapists throughout San Mateo County, San Francisco, and Alameda County; (4) presentations by MHASF staff at Bay Area senior centers, the County of San Mateo Health, Alameda In-home Supportive Services, Hospitality House in San Francisco’s Tenderloin, and the Berkeley City council; (5) flyers distributed at coffee shops, laundromats, libraries, and community centers; (6) announcements made at MHASF’s ongoing drop-in support groups for HD and the Annual Conference on Hoarding and Cluttering; and (7) clinicaltrials.gov.
Study Procedures

Peer facilitators or the MHASF research coordinator conducted an initial screening of interested individuals who contacted the MHASF. The nature of the study, including the randomization, was explained to potential participants, although formal informed consent with a more thorough explanation of all the components of the study was obtained once an individual was enrolled and before randomization. Table 1 outlines the assessments at each time point.

Table 1. Assessments at Pretreatment (Including Screening Questionnaires), Posttreatment, and Longitudinal Follow-up

<table>
<thead>
<tr>
<th>Assessment</th>
<th>Time</th>
<th>Frequency</th>
<th>Timing</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>SI-R</td>
<td>5 min</td>
<td>3 times</td>
<td>Pre-tx, post-tx, longitudinal</td>
<td>Screening, outcome</td>
</tr>
<tr>
<td>UHSS</td>
<td>2 min</td>
<td>Once</td>
<td>Pre-tx</td>
<td>Screening</td>
</tr>
<tr>
<td>CI-R</td>
<td>1 min</td>
<td>Twice</td>
<td>Pre-tx and post-tx</td>
<td>Screening, outcome</td>
</tr>
<tr>
<td>SCI</td>
<td>5 min</td>
<td>Twice</td>
<td>Pre-tx and post-tx</td>
<td>Outcome</td>
</tr>
<tr>
<td>ADL-H</td>
<td>3 min</td>
<td>3 times</td>
<td>Pre-tx, post-tx, longitudinal</td>
<td>Outcome</td>
</tr>
<tr>
<td>SIHD</td>
<td>30 min</td>
<td>Once</td>
<td>Pre-tx</td>
<td>Inclusion</td>
</tr>
<tr>
<td>MINI</td>
<td>15 min</td>
<td>Once</td>
<td>Pre-tx</td>
<td>Inclusion, predictor</td>
</tr>
<tr>
<td>MoCA</td>
<td>5 min</td>
<td>Once</td>
<td>Pre-tx</td>
<td>Exclusion</td>
</tr>
<tr>
<td>BDI</td>
<td>2 min</td>
<td>Twice</td>
<td>Pre-tx and post-tx</td>
<td>Predictor</td>
</tr>
<tr>
<td>BAI</td>
<td>2 min</td>
<td>Twice</td>
<td>Pre-tx and post-tx</td>
<td>Predictor</td>
</tr>
<tr>
<td>Neuropsych battery</td>
<td>60 min</td>
<td>Twice</td>
<td>Pre-tx and post-tx</td>
<td>Predictor, outcome</td>
</tr>
<tr>
<td>Treatment evaluation</td>
<td>2 min</td>
<td>16 times</td>
<td>Weekly during tx</td>
<td>Outcome</td>
</tr>
<tr>
<td>Beliefs assessment</td>
<td>2 min</td>
<td>Twice</td>
<td>Pre-tx and post-tx</td>
<td>Predictor</td>
</tr>
<tr>
<td>Continuing support assessment</td>
<td>2 min</td>
<td>Once</td>
<td>Longitudinal</td>
<td>Predictor of long-term outcome</td>
</tr>
</tbody>
</table>

Abbreviations: ADL-H, Activities of Daily Living Scale for Hoarding Disorder; BAI, Beck Anxiety Inventory; BDI, Beck Depression Inventory; CI-R, Clutter Imaging Rating Scale; MINI, Mini International Neuropsychiatric Interview; MoCA; Montreal Cognitive Assessment; SCI, Saving Cognition Inventory; SIHD, Structured Interview for Hoarding Disorder; tx, treatment; UHSS, UCLA Hoarding Symptom Scale.

The screening instruments were SI-R, the UCLA Hoarding Symptom Scale (UHSS), and the Clutter Image Rating Scale (CI-R). All assessments are described in detail in the Data Collection section below. Individuals aged 18 and older who met the cutoff scores for clinically significant hoarding on 2 of the 3 measures (SI-R total score ≥ 42, UHSS total score ≥ 20, CI-R total score ≥ 12) were contacted by UCSF staff for a clinical interview to confirm a diagnosis of HD using the Structured Instrument for Hoarding Disorder (SIHD). Upon confirmation of
eligibility, we assessed other psychiatric diagnoses using the Mini International Neuropsychiatric Interview (MINI),61 followed by a screening test for dementia using the Montreal Cognitive Assessment (MoCA),62,63 and a battery of neuropsychological tests assessing visual and verbal learning and memory, categorization, information processing speed, visual spatial perception, abstract reasoning abilities, and attention.28

Individuals with moderate to severe dementia symptoms (MoCA ≤ 17) were excluded due to an associated high rate of collecting or hoarding behaviors that differ phenotypically and etiologically from HD. Individuals with current or past psychiatric comorbidities or symptoms were not excluded unless they were at imminent risk for suicide, requiring hospitalization, or were disruptive and could not be redirected during the initial evaluation or first group treatment session. Individuals who had received CBT for hoarding (group or individual) or PFT in the past year were excluded. We chose this time frame because our clinical experience suggested that the gains received from HD-specific treatment would be stable after a year. Individuals who received non-CBT forms of treatment for HD (eg, Clutterers Anonymous, medication) or CBT treatments for nonhoarding disorders were not excluded. However, this information was collected for use in future analyses. Individuals who were excluded or chose not to participate in the study were offered information about the ongoing support and drop-in group programs offered by MHASF.

Eligible participants were randomized to either group CBT or group PFT. Prior to and/or upon completion of the treatment, self-report questionnaires were administered, assessing (1) demographics, (2) HD symptom severity (SI-R, UHSS), (3) functional impairment related to HD (Activities of Daily Living Scale for Hoarding Disorder [ADL-H]), (4) hoarding-related beliefs (Saving Cognition Inventory [SCI]), (5) depression and anxiety symptoms (Beck Depression Inventory-II [BDI], Beck Anxiety Inventory [BAI]), and (6) treatment preference and beliefs (Beliefs About Treatment Questionnaire; see Appendix 2). During treatment, a weekly Homework Completion and Group Treatment Evaluation Form (including information on presence of a clutter buddy) (see Appendix 3) were given to the participants to complete.

Prior to and after treatment, assessments on a number of neuropsychological functions were conducted, which included the Brief Visuospatial Memory Test–Revised (BVMT-R; visual
learning and memory), Hopkins Verbal Learning Test–Revised (HVLT-R; verbal learning and memory), Delis–Kaplan Executive Function System-Sorting Test (categorization), Symbol Digit Modalities Test (SDMT) and Stroop (information processing speed), Wechsler Adult Intelligence Scale-IV Block Design (visual spatial perception), Wechsler Adult Intelligence Scale-IV Matrix (abstract reasoning abilities), Connors Continuous Performance Test II (CPT; attention), Digit Span (memory), and Iowa Gambling Task (IGT; decision making). The National Adult Reading Test (NART) was administered as a proxy for intelligence and included as a covariate in all analyses. As this was an exploratory aim, we deliberately chose a neuropsychological battery to examine a wide range of cognitive domains rather than focusing on a few, in order to identify all potential patterns. This battery took approximately an hour to complete, on average, and was administered by research staff trained in the administration of neuropsychological assessments, overseen by Dr. Mackin, a licensed neuropsychologist.

All participants were financially compensated for their time. At the completion of the pretreatment neuropsychological testing, participants were paid $30. At the completion of the posttreatment neuropsychological testing, participants were paid $70 if they had completed some or all of the treatment groups. If participants completed the posttreatment neuropsychological testing but had not attended any group treatments, they were compensated with $30. Additionally, we offered $30 to those who did not complete the groups or posttreatment neuropsychological testing but were willing to complete the posttreatment questionnaires.

Our efforts to reduce attrition of participants included the following: If a participant began a group, but attended fewer than 3 sessions before dropping out, he or she was eligible to participate in another group of the same type (eg, CBT or PFT) at a later date. We contacted individuals as soon as possible after completing treatment to conduct posttreatment assessments, and the bulk of the compensation for participation was given after the posttreatment assessment was completed. Follow-up contact was made by phone and/or email, and if necessary, by mail. To maximize follow-up, at least 3 attempts to contact participants were made. Group leaders or coordinators asked the participants why they stopped or withdrew. Participants at follow-up were readministered the SI-R and the ADL-H to
participants at follow-up.

**Study Setting**

CBT and PFT groups were run concurrently at 4 locations, 2 in San Francisco, 1 in the East Bay (Berkeley), and 1 in the South Bay (San Mateo) (Table 2). Each location (ie, UCSF, MHASF, South Berkeley Senior Center, and Heart & Soul in San Mateo) provided a private conference room or other similarly sized room for use by the study, with all locations easily accessible by public transportation. Both evening and daytime groups were offered. At any given time during the study treatment period, 6 to 8 groups (3-4 CBT and 3-4 PFT) were conducted per week. Group size was optimized at 8 to 12 members, but for logistical reasons some groups began with fewer in order to minimize delays in starting treatment. We defined dropouts as individuals who completed fewer than 60% of the sessions or those who did not complete the posttreatment assessments.
Table 2. Schedule for CBT and PFT groups. Of the 160 randomized to CBT, 155 attended at least one session (enrolled), while 5 did not enroll. Of the 163 randomized to PFT, 152 attended at least one session, while 11 did not enroll.

<table>
<thead>
<tr>
<th>Round</th>
<th>Day of Week</th>
<th>Facilitator(s)</th>
<th>Location</th>
<th>No. Enrolled</th>
<th>Time</th>
<th>Start Date</th>
<th>No. Graduated</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CBT</td>
<td>PFT</td>
<td>CBT</td>
<td>PFT</td>
</tr>
<tr>
<td>1</td>
<td>Wed</td>
<td>Thurs</td>
<td>Chou/Uhm</td>
<td>10</td>
<td>9</td>
<td>5:30 PM</td>
<td>5/7/14</td>
</tr>
<tr>
<td>2</td>
<td>Tues</td>
<td>Wed</td>
<td>Chou/Uhm</td>
<td>12</td>
<td>9</td>
<td>5:30 PM</td>
<td>6/17/14</td>
</tr>
<tr>
<td>3</td>
<td>Mon</td>
<td>Tues</td>
<td>Uhm</td>
<td>10</td>
<td>8</td>
<td>10:00 AM</td>
<td>8/11/14</td>
</tr>
<tr>
<td>4</td>
<td>Thurs</td>
<td>Mon</td>
<td>Chou</td>
<td>9</td>
<td>13</td>
<td>12:00 PM</td>
<td>9/11/14</td>
</tr>
<tr>
<td>5</td>
<td>Wed</td>
<td>Thurs</td>
<td>Uhm</td>
<td>8</td>
<td>7</td>
<td>5:30 PM</td>
<td>10/15/14</td>
</tr>
<tr>
<td>6</td>
<td>Tues</td>
<td>Wed</td>
<td>Chou</td>
<td>10</td>
<td>8</td>
<td>5:30 PM</td>
<td>12/2/14</td>
</tr>
<tr>
<td>7</td>
<td>Mon</td>
<td>Tues</td>
<td>Uhm</td>
<td>9</td>
<td>8</td>
<td>5:30 PM</td>
<td>2/23/15</td>
</tr>
<tr>
<td>8</td>
<td>Thurs</td>
<td>Mon</td>
<td>Chou</td>
<td>14</td>
<td>13</td>
<td>12:30 PM</td>
<td>2/26/15</td>
</tr>
<tr>
<td>9</td>
<td>Tues</td>
<td>Mon</td>
<td>Uhm</td>
<td>6</td>
<td>10</td>
<td>11:30 AM/10:00 PM</td>
<td>3/17/15</td>
</tr>
<tr>
<td>10</td>
<td>Wed</td>
<td>Thurs</td>
<td>Chou</td>
<td>12</td>
<td>6</td>
<td>5:30 PM</td>
<td>4/8/15</td>
</tr>
<tr>
<td>11</td>
<td>Thurs</td>
<td>Wed</td>
<td>Uhm</td>
<td>9</td>
<td>6</td>
<td>5:30 PM</td>
<td>5/28/15</td>
</tr>
<tr>
<td>12</td>
<td>Thurs</td>
<td>Mon</td>
<td>Chou</td>
<td>8</td>
<td>9</td>
<td>12:30 PM</td>
<td>7/30/15</td>
</tr>
<tr>
<td>13</td>
<td>Tues</td>
<td>Mon</td>
<td>Uhm</td>
<td>6</td>
<td>6</td>
<td>10:30 AM/12:30 PM</td>
<td>8/18/15</td>
</tr>
<tr>
<td>14</td>
<td>Mon</td>
<td>Tues</td>
<td>Chou</td>
<td>5</td>
<td>6</td>
<td>5:30 PM</td>
<td>9/28/15</td>
</tr>
<tr>
<td>15</td>
<td>Wed</td>
<td>Thurs</td>
<td>Uhm/Chou</td>
<td>5</td>
<td>8</td>
<td>6:30 PM/5:30 PM</td>
<td>10/28/15</td>
</tr>
<tr>
<td>16</td>
<td>Thurs</td>
<td>Mon</td>
<td>Chou</td>
<td>9</td>
<td>11</td>
<td>12:30 PM</td>
<td>1/21/16</td>
</tr>
<tr>
<td>17</td>
<td>Thurs</td>
<td>Wed</td>
<td>Chou</td>
<td>13</td>
<td>15</td>
<td>5:30 PM</td>
<td>3/3/16</td>
</tr>
</tbody>
</table>
Interventions

The interventions in this study included group CBT using an HD-specific CBT treatment manual and facilitators’ guide and group PFT using the BiT self-help workbook and associated peer facilitators’ guide. Group treatment was chosen over individual to allow for the provision of treatment to the largest number of participants possible. The target number of participants in each group was 8 to 12, the standard size of the groups run by MHASF. Although this number is larger than seen in other studies, it was considered to be feasible by the MHASF treatment providers with experience and allowed for possible attrition of group members.

CBT groups: Based on the experience of MHASF in running CBT groups with community-based clinicians as leaders, and to make the CBT groups more comparable to the PFT groups in terms of number of sessions, we adapted the manual-based treatment to be 16 rather than 20 sessions, conducted over 20 weeks. We also assigned “clutter buddies” from within the group rather than recommending that participants find their own clutter buddies. Finally, we included 2 home visits as outlined in the manual, although these visits were 30 minutes rather than an hour in length. The first visit occurred after the third session rather than prior to treatment to decrease fears of stigma and judgment, based on MHASF’s experience that participants were reluctant to let providers into their homes prior to developing a relationship with them. The second visit occurred after the 15th session. The aim of the visits was to objectively assess the degree of clutter in participants’ homes.

As detailed in the CBT therapists’ guide (and in the PFT facilitators’ guide, used in the PFT component as described below), each 2-hour session consisted of check-ins and psychoeducation to accomplish the objectives, including understanding and awareness of one’s hoarding symptoms and patterns, goal setting, behavior modification, cognitive restructuring, motivational enhancement, in vivo and imagery exposure, executive skills training, relapse prevention, and establishing clutter buddies. In the CBT treatment groups, the leaders assigned each participant a clutter buddy from within his or her group, typically based on geographic location or other logistical considerations, although individual requests for a specific clutter buddy assignment were honored. The purpose of the clutter buddies was to provide motivation
and help with adherence to assigned homework, as well as overall support. Participants were asked in the weekly evaluations whether and how frequently they met with their clutter buddies during the previous week. Weekly homework assignments included reading and practicing skills learned in sessions.

**PFT groups:** The PFT groups consisted of 15 sessions over 20 weeks led by 1 to 2 trained peers with personal lived experience of hoarding using the treatment manual *Buried in Treasures: Help for Compulsive Acquiring, Saving, and Hoarding.*

Peer facilitators in the PFT groups chose not to assign clutter buddies but rather to encourage the group participants to identify a clutter buddy from friends, family, or other group members. The topics of psychoeducation in the PFT group were similar to those in the CBT group (see Appendix 1). As with the CBT group, each PFT group was 2 hours in length with homework assigned weekly. Additionally, peer facilitators called each member of the group weekly prior to the group meeting to check in and field questions, if any. The PFT treatment did not include home visits.

**Supervision and adherence to treatment manuals:** Group facilitators for both CBT and PFT received initial training in delivering the treatment (by Dr. Joanne Chan and Mr. Lee Shuer, respectively) in a 2-day training workshop. They then received an hour of supervision each week (by Drs. Chan and Mathews for CBT and by Dr. Eckfield and Mr. Plumadore for PFT). Although adherence to the treatment manual was not formally assessed (eg, through videotaping and reviewing treatment sessions), adherence was assessed during the supervision sessions in an informal way by discussing what occurred during the treatment sessions with the facilitators and providing directions regarding how to facilitate upcoming sessions. The PFT training workshop focused on guiding participants through the BiT workbook and provided rudimentary training on facilitating groups. The CBT training focused on working with the therapists’ guide and familiarizing the postdoctoral fellows with the accompanying workbook.

**Follow-up**

The scheduled assessments included eligibility screening, pretreatment and posttreatment (within a month after treatment) assessment on clinical and neuropsychological
function, homework compliance, and treatment satisfaction (assessed during treatment period). While longitudinal follow-up assessments were not a part of the original protocol, in the final year of the study, in discussion with the PCORI project officer, we obtained additional funding that allowed us to recontact participants at least 3 months (range 3-25 months) after they had completed treatment and conduct longitudinal assessments to assess any maintenance (or loss) of treatment gains. For this component we contacted all participants who were randomized for treatment and had given us permission to recontact them for possible participation in future studies; 6 participants did not give us permission for recontact and thus were not included. Participants who agreed to this additional assessment signed a new consent form and were paid an additional $20 for their time. As we decided to conduct this long-term follow-up after the study began, and study was in its final year, the elapsed time from end of treatment to longitudinal assessment varied. Participants who had not yet completed treatment were consented prospectively for this component of the research and reassessed at 3 months posttreatment. The longitudinal assessment, which was conducted by a member of the research team who was blind to treatment arm and posttreatment outcome, consisted of the 2 primary outcome measures of interest, the SI-R and the ADL-H, as well as questions about the participants’ ongoing involvement in hoarding-related supports or treatments, if any (see Appendix 4).

Study Outcomes

The primary outcome variable was posttreatment score on the SI-R, and the primary predictor variable was group status (CBT versus PFT). The secondary outcome measure included posttreatment scores on activities of daily living/daily functioning measured by the ADL-H. Additional exploratory secondary outcomes included treatment adherence and neurocognitive status posttreatment.

Data Collection and Sources

All participants provided informed consent to participate in the research and in all study procedures. The main study protocol was as follows: Participants completed self-report questionnaires online, by telephone, or in person as an initial screen for hoarding symptoms. Those who screened positive for significant HD symptoms (through scores on the SI-R, UHSS,
and CI-R) provided informed consent to participate in the remaining components of the study, and completed additional self-report questionnaires assessing depression, anxiety, attentional, and other psychiatric symptoms relevant to HD (SCI, BDI, BAI, and Swanson Nolan and Pelham Rating Scale for ADHD). They then participated in a clinical interview (SIHD) aimed at assessing HD (based on DSM-5), other cooccurring psychiatric disorders (MINI, based on DSM-IV-TR), and symptoms of dementia (MoCA), and underwent formal neuropsychological assessment. All assessments are detailed below.

**Saving Inventory-Revised**: We used the SI-R both as the primary outcome measure and as 1 of 3 initial screening measures. The SI-R is a 23-item self-report questionnaire that measures hoarding symptoms and their impact, including problems with acquisition, clutter, and difficulty discarding, as well as distress and impairment/interference. Each question is scored on a 4-point Likert scale ranging from 0 (none or no symptoms) to 4 (very much or extreme); the total possible score is 96, with a higher score indicating more severe hoarding symptoms. The SI-R has 3 subscales (difficulty discarding, acquisition, and clutter). Mean scores for help-seeking individuals with hoarding are generally in the 60s. The SI-R is widely used in hoarding research, has good test–retest reliability (kappa = 0.86), internal consistency (Cronbach \( \alpha \) between 0.91 and 0.94), and reliably discriminates between HD and community controls. A cutoff score of 41 has been suggested as optimal to identify HD based on receiver operating characteristic graphs; our inclusion criteria required an SI-R score of > 41. Most treatment studies of HD report baseline SI-R scores in the mid-60s for HD. Individuals without HD typically have scores in the 20s. We used a change score (improvement) of 14 points or more as the criterion for clinically significant improvement based on previous studies using this measure as the outcome, while we considered a change score of 10 points or more clinically relevant (noticeable and meaningful improvement from a patient’s and health care provider’s perspective) (pers comm, R. Frost and based on the average improvement seen in our pilot study).

**Activities of Daily Living Scale for Hoarding Disorder**: We used the ADL-H as the secondary outcome measure, and to assess functional impairment due to hoarding symptoms. The ADL-H
is a self-report measure containing 15 items that are scored on a 5-point scale ("can do it easily" to "unable to do"). Although the ADL-H items can be summed, resulting in a total possible score of 75, the mean score of the individual items is more typically used, to account for items that are not relevant to a particular individual (eg, clutter on stairs). Mean scores for individuals with hoarding range between 1.95 and 2.2; community controls have mean scores of about 1.15. The ADL-H has good internal reliability (alpha = 0.96), test–retest reliability ($r = 0.79$), and interrater reliability ($r = 0.71$). Previous studies have found changes of about 0.5 points (approximately 25%) using the item mean scoring method following treatment for hoarding.

_UCLA Hoarding Severity Scale_\(^{52}\): The UHSS is a 10-item clinician-administered instrument that was designed to be used in conjunction with a clinician interview. The UHSS uses a Likert scale scoring system from 0 to 4 for each question and assesses clutter, acquisition, and difficulty discarding, as well as the individual’s level of shame associated with hoarding behaviors, his or her impairment in social relationships due to hoarding, procrastination, and difficulty making decisions. The authors have proposed a cutoff score of > 20 to indicate clinically significant compulsive hoarding.\(^{52}\) The UHSS has good internal consistency (Cronbach $\alpha = 0.70$).\(^{70}\) We used the UHSS as a self-report instrument to assess for hoarding symptoms prior to entry into the study.

_Clutter Image Rating Scale_\(^{71}\): The CI-R uses a series of 9 photographs each of a kitchen, living room, and bedroom depicting varying levels of clutter, as an adjunct to visually measure the severity of clutter. Participants select the photograph that most closely resembles each of the 3 rooms in their home. Each room is scored on a scale of 1 to 9, for a total possible score of 27. Scores of $\geq 4$ on each of the 3 rooms are typically considered clinically significant.\(^{71}\) We used a score of $> 12$ on the CI-R as the cutoff in the screening portion of the study, as we have done in previous studies.\(^{72,73}\) Internal consistency ($\alpha = .84$), test–retest reliability ($r = 0.82$), and interrater reliability ($r = 0.94$) for the CI-R are high.\(^{71}\) For individuals who lived in a single room, their CI-R score for the room that they rated was tripled.

_Structured Interview for Hoarding Disorder_: We used the SIHD to confirm a diagnosis of HD.\(^{6,60}\)
The SIHD is a semistructured clinical interview instrument designed to assess for HD according to the proposed DSM-5 criteria. The SIHD assesses hoarding behaviors and distress and interference associated with hoarding behaviors and/or with discarding items, as well as assesses for potential confounders or alternate diagnoses, such as autism or schizophrenia.

Mini International Neuropsychiatric Interview\textsuperscript{61}: We used the MINI to assess presence of lifetime and current history of psychiatric disorders prior to randomization. The MINI assesses for presence of all major psychiatric disorders (eg, mood, anxiety, psychotic, substance use disorders).\textsuperscript{61} The MINI has good reliability and validity and is widely used in neuropsychiatric research as a brief diagnostic screen.

Saving Cognitions Inventory: We assessed hoarding beliefs using the SCI.\textsuperscript{74} The SCI is a 24-item scale assessing 4 types of beliefs and attitudes about possessions (emotional attachment, need to control, responsibility, memory). Responses range from 1 (not at all agree) to 7 (very much agree), with higher scores indicating stronger hoarding-related beliefs. The total possible score is 168. The SCI has high internal consistency, and convergent and discriminant validity.\textsuperscript{74,75} Average scores for individuals with HD are about 100\textsuperscript{74}; improvement in SCI scores has been shown to mediate improvement in hoarding symptom severity following treatment.\textsuperscript{76}

Beck Depression Inventory, Second Edition and Beck Anxiety Inventory \textsuperscript{77,78}: We used the BDI-II and BAI to assess current depression and anxiety symptoms both pretreatment and posttreatment. Both the BDI-II and the BAI are 21-item self-report questionnaires. The BDI was designed to measure severity of current depressive symptoms, and the BAI was designed to measure severity of current anxiety symptoms. Items are scored on a 0-to-3 scale. The total possible score on each measure is 63. Higher scores indicate more severe depressive or anxious symptoms. We included the total score on each of these measures in the analyses to control for the influence of concurrent depressive and/or anxiety symptoms on the outcomes.

The Swanson, Nolan and Pelham Questionnaire (SNAP-IV)\textsuperscript{79}: We used the SNAP-IV to assess symptoms of attention, concentration, hyperactivity, and impulsiveness. The SNAP-IV consists of 10 inattentive and 10 hyperactive-impulsive symptoms, all rated on a 4-point scale, ranging
from not at all to very much. Higher scores indicate more severe symptoms. The SNAP-IV total score has an area under the curve of 0.90, with an optimal cutoff score of 24.5 to identify ADHD.\textsuperscript{80} We summed the SNAP-IV and included the total score in the analyses to assess the influence of ADHD symptoms on the outcomes.

The Montreal Cognitive Assessment\textsuperscript{63,81}: We used the MoCA to assess for symptoms of cognitive impairment and to exclude individuals with dementia. The MoCA was administered by the clinicians who conducted the clinical interviews. We excluded individuals who scored \( \geq 17 \) of a total possible 30 (considered moderate to severe dementia).

Beliefs About Treatment Questionnaire (Appendix 2): This questionnaire was designed by the stakeholders at MHASF and was intended to assess, in a quantitative manner, participants’ beliefs about how beneficial group therapy for hoarding would be (or was, posttreatment) in general, and in particular, how beneficial they would find (or found) PFT or CBT. There are no validity or reliability data available for this instrument, as it was designed for this study.

Continuing Support Questionnaire (Appendix 4): This questionnaire was designed by the MHASF research team for the longitudinal assessment component and asks about types of ongoing support or treatment that the participants may have continued after the formal end of the treatment study. There are no validity or reliability data available for this measure.

**Analytical and Statistical Approaches**

We conducted all analyses using SAS v9.4. We used standard methods to summarize and describe the collected data, including means, standard deviations, frequencies, rates, and patterns of participant attrition. A tally of the number who met criteria for clinically meaningful change (SI-R change \( \geq 10 \) points), clinically significant change (SI-R change \( \geq 14 \) points),\textsuperscript{39,69} and functional remission defined as a clinically significant reduction in SI-R, eg, at least 14 points and to below the previously suggested clinical cutoff for HD(SI-R < 42)\textsuperscript{39,67,69} was generated. We conducted all analyses (including the longitudinal follow-up) under the intent-to-treat principle, with participants classified to the condition they were randomized to regardless of the extent of any postrandomization treatment.
To test for noninferiority of treatment conditions (aim 1), we used a 1-tailed $t$ test for noninferiority comparing the mean posttreatment SI-R scores with a margin of equivalence equal to 5. This margin is deliberately conservative, as it represents half of the minimal change score that was determined to be clinically meaningful (SI-R change $\geq 10$ points, or 15% improvement from a mean SI-R baseline of 65). As there are no published data regarding meaningful change scores for the ADL-H, we arbitrarily set the margin of 2.5 using a similar approach (eg, minimum of 15% improvement from baseline total score divided by 2) in the secondary exploratory analyses. We conducted testing for individual characteristics that were associated with treatment response by the estimation and testing of a statistical regression model of change in SI-R scores, which incorporated each individual measure (eg, insurance status and socioeconomic, demographic, clinical, and neuropsychological factors) as a covariate. We conducted these univariable analyses across the entire data set and across both treatment arms. Participant preferences for elements incorporated in CBT or PFT were collected prior to randomization at baseline, measured on a scale of 1 to 5 (strong preference against to strong preference for each treatment), and examined for their association with the treatment outcome by correlating preference (measured as strong preference for CBT or PFT or no preference) with the outcome measure. We computed the effect size of the change on the 2 primary outcomes as mean change divided by the standard error of the change. We examined the change in the 2 main outcomes (SI-R and ADL-H) by treatment group over time (pretreatment, posttreatment, longitudinal) using a mixed linear model (the mixed procedure in SAS) that incorporated all 3 time points. As an exploratory analysis, we also examined predictors of response ($\geq 10$-point improvement on the SI-R), no change (4- to 9-point change on the SI-R) and worsening of symptoms ($\geq -5$-point change on the SI-R) at posttreatment and longitudinal follow-up.

Missing data: We constantly monitored collected data to check for missing values in questionnaires. When found, we made every attempt to recover the data by checking for hard copies, contacting participants if available, resending surveys to participants, or requesting relevant information (eg, regarding group participation). For individual data points, we compared baseline and posttreatment missingness rates between participants who provided
posttreatment data and those who did not using the appropriate test (t test or Pearson chi-square) to test for evidence that the missing data were not missing at random. In this pre–post design, the likelihood for the outcome given treatment condition is empty when the outcome is missing. As the missing outcomes cannot be meaningfully imputed, we included only participants with complete outcome data for any given measure in the analyses. Thus, the numbers of participants varied in the secondary analyses based on whether the outcome of interest was available. We also assessed demographic and clinical characteristics of participants who completed baseline but not posttreatment data.

Sensitivity analyses: To test for evidence of effects of participant dropout, we reestimated and tested our models of noninferiority and predictors of treatment response using only participants who completed treatment (completers-only analysis). In a similar way, we conducted sensitivity analyses to examine the effects of including or excluding participants who dropped out of treatment prior to the third session and later rejoined a group (N = 12). We note that these analyses should be treated as exploratory and the results interpreted with caution.

Power calculations: We chose the sample size for this study (N = 300 total, N = 150 per arm) using power calculations conducted prior to the initiation of the study, and based on data showing a mean improvement in the SI-R of about 10 points in our pilot work,49 and increased the sample size screened and assessed to account for attrition, to ensure that we were able to randomize at least 150 individuals in each treatment arm. Our power calculations, which we conducted during the study design phase, suggested that with a sample size of 150 participants per arm, we had 80% power to demonstrate noninferiority if the PFT mean was no more than 2.9 points below (ie, worse than) the CBT mean. We based our power calculations on the assumption that a 3-point difference between the 2 groups on a scale where the mean score for participants entering the study was approximately 65, and improvement was anticipated to be between 10 and 14 points, was unlikely to be clinically meaningful, and therefore, any difference between the groups that was smaller than 3 points would signify that the 2 treatment types were equally effective.
Conduct of the Study

The study was approved by the UCSF Institutional Review Board. As previously described, the final study protocol was very similar to what we originally proposed. We made minor changes to facilitate the longitudinal assessment, including modifying the consent form to ask for participants’ permission to be recontacted by the research team at UCSF and MHASF for future research and clinical services (CHR reference number: 136105), recontacting participants for the longitudinal component to assess treatment response 3 or more months following treatment (CHR reference number: 172226), and updating the study site to reflect Dr. Mathews’ move to the University of Florida during the second year of the study (CHR reference number: 168542).
Results

Six hundred and thirty-two individuals expressed interest in the study (Figure 1). Of these, 476 completed the screening process, and 414 were clinically assessed. Of the 476 individuals who were screened, 13% (N = 62) were excluded based on inclusion/exclusion criteria; the remainder (N = 156) either declined to participate (N = 53), did not complete the screen (N = 68), or were lost to follow-up (N = 35). Only 3% of the 414 individuals who underwent clinical assessment were found to be ineligible for participation (N = 12); the remainder, who were not randomized for treatment, either declined to participate or were lost to follow-up. Of the 414, 323 were randomized into CBT or PFT. Two hundred thirty-one individuals completed treatment (69.3% in PFT and 73.7% in CBT). Reasons for noncompletion are documented in Figure 1 (some individuals had more than one reason). Of the 323 who were randomized, 247 completed the posttreatment assessment, and of the 317 contacted for the longitudinal data who consented to recontact (6 individuals did not consent to recontact), 183 completed it. Analysis of missing data indicated that if the participant completed the assessment in question, essentially no individual item data were missing (mean numbers of missing items were 0.08 and 0.03 for the for the SI-R and ADL-H, respectively). There were no differences between completion rates or missingness by treatment arm ($\chi^2 = 0.14$, $df = 1$, $p = 0.71$ for SI-R; $\chi^2 = 0.82$, $df = 1$, $p = 0.36$ for ADL-H), questionnaire, or time point (pretreatment versus posttreatment versus longitudinal).

Individuals who completed the screening but were not randomized for treatment (N = 149) were more likely to be men, have poor insight into their illness, and have lower mean hoarding severity scores but similar hoarding-related impairment rates than were individuals who were randomized (Table 3). Race and ethnicity were not assessed at screening.
Table 3. Characteristics of individuals completing the pretreatment screening. Poor insight was assessed by clinical interview in 334 participants (323 who were randomized to treatment and 11 who were excluded or chose not to participate).

<table>
<thead>
<tr>
<th></th>
<th>Screened, Not Randomized (N = 149)</th>
<th>Randomized (N = 323)</th>
<th>t Statistic (df = 1)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>SI-R total score</td>
<td>54.9 (18.3)</td>
<td>65.5 (11.7)</td>
<td>−7.59</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>UHSS total score</td>
<td>23.2 (8.6)</td>
<td>29.1 (5.5)</td>
<td>−8.88</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>CI-R total score</td>
<td>10.6 (5.0)</td>
<td>13.5 (5.0)</td>
<td>−5.71</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>ADL-H total</td>
<td>29.5 (11.0)</td>
<td>31.0 (10.5)</td>
<td>−0.55</td>
<td>0.58</td>
</tr>
<tr>
<td>Poor insight</td>
<td>3/11 (27.3%)</td>
<td>39/323</td>
<td>Fisher exact</td>
<td>0.009</td>
</tr>
<tr>
<td>Age a</td>
<td>57.6 (13.3)</td>
<td>59.1 (10.6)</td>
<td>−1.37</td>
<td>0.17</td>
</tr>
<tr>
<td>Female a</td>
<td>138 (61.9%)</td>
<td>241 (74.6%)</td>
<td>10.07 (df = 2)</td>
<td>0.007</td>
</tr>
</tbody>
</table>

Abbreviations: ADL-H, Activities of Daily Living Scale, Hoarding; CI-R, Clutter Image Rating Scale; SI-R, Saving Inventory, Revised; UHSS, UCLA Hoarding Severity Scale.

a Data available for 223 of the nonrandomized participants.
Figure 1. Study Flow Diagram
Baseline (Pretreatment) Characteristics

With the exception of UHSS total score, which was, on average, 1 point higher for PFT than for CBT participants, there were no significant differences in key baseline clinical or demographic characteristics between participants randomized to CBT and those randomized to PFT (Table 4). The mean age of the sample was approximately 59 years and 74.6% were female. Of the 328 for whom we have data, 59.5% of participants reported their race as white, 11.3% reported being Asian, 7.6% reported being African American, 1.2% reported being Native American or Native Hawaiian/Pacific Islander, 20.4% reported either having more than one race or “other race,” and 9.3% reported being Hispanic. Most participants were insured and had at least a college education. Most were not married and lived alone, and fewer than one-third were employed at the time of participation.

Participants had moderately severe hoarding symptoms at baseline as measured by the SI-R and by the UHSS (Table 4); 61% had a current or lifetime history of a psychiatric disorder other than HD (34% had 2 or more disorders), and as a group, participants reported moderate levels of depression and anxiety (measured by the BDI and BAI, respectively). Of participants, 49.8% had a mood disorder, 48.2% had an anxiety disorder, 2.8% had ADHD, and 5.7% had OCD. Approximately 10% had a moderate to high suicide risk at the time of entry into the study.

Of participants, 72% expressed no strong preference for either treatment prior to randomization. Only 1 of the 10 participants who preferred PFT was randomized to CBT, while 36 of the 77 who preferred CBT were randomized to PFT ($X^2 = 6.72; df = 2; p = 0.04$).
### Table 4. Pretreatment Participant Characteristics

<table>
<thead>
<tr>
<th></th>
<th>CBT (N = 160) Mean (SD) or N (%)</th>
<th>PFT (N = 163) Mean (SD) or N (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex (male)</td>
<td>43 (26.9%)</td>
<td>39 (23.9%)</td>
</tr>
<tr>
<td>Age (mean, (SD))</td>
<td>59.0 (10.9)</td>
<td>58.9 (10.6)</td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>89 (55.6%)</td>
<td>101 (62.4%)</td>
</tr>
<tr>
<td>Asian</td>
<td>16 (10.0%)</td>
<td>17 (10.5%)</td>
</tr>
<tr>
<td>African American</td>
<td>17 (10.6%)</td>
<td>6 (3.7%)</td>
</tr>
<tr>
<td>Multiracial/other</td>
<td>23 (14.4%)</td>
<td>23 (14.2%)</td>
</tr>
<tr>
<td>Hispanic/Latino</td>
<td>9 (5.6%)</td>
<td>8 (4.9%)</td>
</tr>
<tr>
<td>LGBT</td>
<td>22 (15.8%)</td>
<td>22 (16.3%)</td>
</tr>
<tr>
<td>Marital status</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Married</td>
<td>40 (26.5%)</td>
<td>37 (24.2%)</td>
</tr>
<tr>
<td>Divorced/widowed/separated</td>
<td>55 (36.4%)</td>
<td>48 (31.4%)</td>
</tr>
<tr>
<td>Never married/roommate</td>
<td>56 (37.1%)</td>
<td>68 (44.4%)</td>
</tr>
<tr>
<td>Number of people in household (adults and children)</td>
<td>1.13 (0.74)</td>
<td>1.21</td>
</tr>
<tr>
<td>Children (Y/N)</td>
<td>88 (56.4%)</td>
<td>0.69</td>
</tr>
<tr>
<td>Employment status</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Employed (full or part time)</td>
<td>54 (34.4%)</td>
<td>29 (25.0%)</td>
</tr>
<tr>
<td>Unemployed/disabled</td>
<td>55 (35.0%)</td>
<td>65 (41.7%)</td>
</tr>
<tr>
<td>Retired</td>
<td>37 (23.6%)</td>
<td>40 (25.6%)</td>
</tr>
<tr>
<td>Other</td>
<td>11 (7.0%)</td>
<td>12 (7.7%)</td>
</tr>
<tr>
<td>Insurance status (insured versus under or not insured)</td>
<td>150 (95.5%)</td>
<td>144 (90.0%)</td>
</tr>
<tr>
<td>Education</td>
<td>15.3 (2.2)</td>
<td>15.3 (2.4)</td>
</tr>
<tr>
<td>Pretreatment MoCA score</td>
<td>26.3 (2.9)</td>
<td>26.2 (2.6)</td>
</tr>
<tr>
<td>Pretreatment NART</td>
<td>115.6 (6.9)</td>
<td>116.0 (6.1)</td>
</tr>
<tr>
<td>Pretreatment UHSS score</td>
<td>28.5 (5.5)</td>
<td>29.8 (5.5)</td>
</tr>
<tr>
<td>Pretreatment SI-R score</td>
<td>64.5 (11.7)</td>
<td>66.4 (11.6)</td>
</tr>
<tr>
<td>Pretreatment ADL-H score</td>
<td>29.9 (9.5)</td>
<td>32.0 (11.2)</td>
</tr>
<tr>
<td>Pretreatment CI-R score</td>
<td>13.0 (4.8)</td>
<td>14.0 (5.1)</td>
</tr>
<tr>
<td>Pretreatment SCI score</td>
<td>97.8 (29.7)</td>
<td>100.2 (29.1)</td>
</tr>
<tr>
<td>Pretreatment BAI score</td>
<td>17.6 (12.5)</td>
<td>18.6 (12.2)</td>
</tr>
<tr>
<td>Pretreatment BDI score</td>
<td>20.1 (12.2)</td>
<td>19.4 (13.1)</td>
</tr>
<tr>
<td>SNAP-IV (ADHD) total score</td>
<td>38.1 (15.0)</td>
<td>36.1 (13.5)</td>
</tr>
<tr>
<td>Excessive acquisition</td>
<td>98 (64.5%)</td>
<td>87 (60.8%)</td>
</tr>
<tr>
<td>Poor insight</td>
<td>18 (11.3%)</td>
<td>21 (12.9%)</td>
</tr>
<tr>
<td>Age of symptom onset &lt;18 vs ≥18</td>
<td>69 (45.1%)</td>
<td>69 (43.7%)</td>
</tr>
<tr>
<td>Any current or lifetime psychiatric diagnosis (not including HD)</td>
<td>95 (63.1%)</td>
<td>89 (61.0%)</td>
</tr>
<tr>
<td>Mean number of psychiatric diagnoses (not including HD)</td>
<td>0.12 (0.14)</td>
<td>0.10 (0.12)</td>
</tr>
<tr>
<td>Taking psychiatric medications</td>
<td>41.8%</td>
<td>47.5%</td>
</tr>
<tr>
<td>Moderate or high suicide risk</td>
<td>12.1%</td>
<td>8.4%</td>
</tr>
</tbody>
</table>

Abbreviations: ADL-H, Activities of Daily Living Scale, Hoarding; BAI, Beck Anxiety Inventory; BDI, Beck Depression Inventory; CI-R, Clutter Image Rating Scale; MoCA, Montreal Cognitive Assessment; NART, National Adult Reading Test; SCI, Saving Cognition Inventory; SI-R, Saving Inventory- Revised; SNAP-IV, Swanson, Nolan and Pelham ADHD
**Participation in Treatment**

Of the 323 individuals randomized to treatment, 269 (83%) attended at least one session, and 231 (71.5%; N = 118 for CBT, N = 113 for PFT) completed treatment (ie, attended 13 of 16 CBT or 12 of 15 PFT sessions). The mean number of sessions attended was 13.5 (SD = 3.2) for CBT and 11.5 (SD = 3.1) for PFT. There were no significant differences in dropout rates between CBT and PFT (Figure 1). Twelve participants dropped out of their original group within the first 3 sessions and reentered treatment. CBT participants attended more group sessions (73.3% of groups, on average; SD = 34.0) than did PFT participants (57.7% of groups, on average; SD = 34.6) ($t = -4.09; df = 1; p < 0.0001$). Homework completion was similar between groups (54.3% of homework assignments completed for CBT, SD = 32.5; 48.9% for PFT, SD = 33.3) ($t = -1.50; df = 1; p = 0.134$) (Table 5). Although clutter buddies were not required for either treatment arm, they were assigned from within the group in CBT, while participants in PFT were encouraged to find a clutter buddy but were not assigned one. Sixty-nine percent of participants in CBT had a clutter buddy (N = 111), while 37% of participants in PFT had a clutter buddy (N = 61) ($X^2 = 33.11; df = 1; p < 0.0001$). Of those with a clutter buddy, the majority in both groups (82.5% of CBT, 83.8% of PFT) found them to be helpful ($X^2 = 0.03; df = 1; p = 0.86$).

**Table 5.** Treatment Outcomes Immediately Posttreatment and at Longitudinal Follow-up (≥ 3 Months Posttreatment)

<table>
<thead>
<tr>
<th></th>
<th>CBT (N = 128)</th>
<th>PFT (N = 119)</th>
<th>$X^2$ or $t$ Statistic</th>
<th>$P$ Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attendance rate</td>
<td>73.3%</td>
<td>57.7%</td>
<td>-4.09</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Homework</td>
<td>54.3%</td>
<td>48.9%</td>
<td>-1.50</td>
<td>0.13</td>
</tr>
<tr>
<td>Clutter buddy</td>
<td>69%</td>
<td>37%</td>
<td>33.11</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Posttreatment SI-R</td>
<td>45.9 (15.0)</td>
<td>47.8 (14.2)</td>
<td>-1.71</td>
<td>0.04</td>
</tr>
<tr>
<td>Posttreatment ADL-H</td>
<td>25.5 (10.1)</td>
<td>26.1 (8.7)</td>
<td>-1.52</td>
<td>0.05</td>
</tr>
<tr>
<td>Posttreatment SCI</td>
<td>81.1 (30.1)</td>
<td>79.6 (27.8)</td>
<td>-0.40</td>
<td>0.69</td>
</tr>
<tr>
<td>Long f/u SI-R</td>
<td>48.2 (14.9)</td>
<td>47.9 (13.9)</td>
<td>-0.15</td>
<td>0.88</td>
</tr>
<tr>
<td>Long f/u ADL-H</td>
<td>27.6 (11.6)</td>
<td>27.8 (9.7)</td>
<td>0.09</td>
<td>0.93</td>
</tr>
</tbody>
</table>

Abbreviations: ADL-H, Activities of Daily Living Scale, Hoarding; SCI, Saving Cognition Inventory; SI-R, Saving Inventory, Revised.
Primary Treatment Outcomes

**Reduction in hoarding severity (total sample).** Posttreatment data were available for 247 individuals (231 individuals who completed treatment and 16 noncompleters). We did not impute missing data, given the simple pre–post design and minimal evidence of missing data for individual items from the outcome measures. We used all available data in each analysis. There was no indication that the missing data added bias to the findings. The mean posttreatment SI-R score for the entire randomized sample, regardless of treatment completion, was 46.8 (SD = 14.6). The mean change in SI-R total score pretreatment to posttreatment was 17.8 points (SD = 14.7), equivalent to a 26.7% (SD = 21.5) reduction in symptom severity. Of participants, 66% responded to treatment (≥10-point reduction in SI-R scores). Of participants, 37% had posttreatment SI-R scores of < 42. Of participants, 31.2% met criteria for remission (SI-R posttreatment scores of < 42 plus a ≥14-point improvement in SI-R scores pretreatment to posttreatment). Effect size for the change in the SI-R was 1.20 for PFT and 1.21 for CBT.

In the re-estimation and testing using only participants who completed treatment (PFT N = 113, CBT N = 118), we found similar results, indicating dropout did not bias the findings. The same was true when excluding participants who dropped out prior to the third session and rejoined a treatment group later (N = 12), and when using mean change in SI-R/ADL-H scores rather than posttest scores.

**CBT versus PFT.** Baseline data for the SI-R were not significantly different between CBT (mean pretreatment SI-R = 64.5; SD = 11.7) and PFT (mean pretreatment SI-R = 66.4; SD = 11.6) (Table 4). Participants in CBT had a 27.7% reduction in SI-R scores (mean posttreatment SI-R = 45.9; SD = 15.0), while participants in PFT had a 25.6% reduction in SI-R scores (mean posttreatment SI-R = 47.8; SD = 14.2) (Table 6). The test of noninferiority indicated that the hypothesis under test (ie, that PFT outcomes would be statistically worse than CBT outcomes) was rejected (between-group difference = 1.82 points; 95% CI: inf., 4.89; t = −1.71; df = 1; p = 0.04), indicating no significant difference in outcomes between participants in CBT and PFT. The results did not change when the analyses were limited to participants who completed treatment. Of participants, 37% in CBT and 36% in PFT had posttreatment SI-R scores of < 42. Sixty-five
percent of CBT participants and 68% of PFT participants had \( \geq 10 \)-point improvement in SI-R scores, while 55% and 57%, respectively, had \( \geq 14 \)-point improvement in SI-R scores pretreatment to posttreatment. There was no significant difference in remission rates between treatment arms (32.8% for CBT, 29.4% for PFT; \( X^2 = 0.33; df = 1; p = 0.56 \)).
Table 6. Pre-treatment, post-treatment, and longitudinal scores for CBT and PFT groups. All numbers are for all participants who were randomized to treatment with one exception: The SI-R results are reported both for all participants and for only those who completed treatment.

<table>
<thead>
<tr>
<th>Measure</th>
<th>CBT</th>
<th>PFT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean Pre-treatment Total Score (SD)</td>
<td>Mean Pre-treatment Total Score (SD)</td>
</tr>
<tr>
<td>SI-R</td>
<td>64.5 (11.7)</td>
<td>66.4 (11.6)</td>
</tr>
<tr>
<td></td>
<td>45.9 (15.0)</td>
<td>47.8 (14.2)</td>
</tr>
<tr>
<td>SI-R (complete rs only)</td>
<td>64.3 (12.0)</td>
<td>65.6 (11.3)</td>
</tr>
<tr>
<td></td>
<td>45.1 (14.4)</td>
<td>47.2 (14.2)</td>
</tr>
<tr>
<td>ADL-H</td>
<td>29.9 (9.5)</td>
<td>32.0 (11.2)</td>
</tr>
<tr>
<td></td>
<td>25.5 (10.1)</td>
<td>26.1 (8.7)</td>
</tr>
<tr>
<td>CI-R</td>
<td>13.0 (4.8)</td>
<td>14.0 (5.1)</td>
</tr>
<tr>
<td></td>
<td>9.8 (4.4)</td>
<td>10.5 (4.9)</td>
</tr>
<tr>
<td>SCI</td>
<td>97.8 (29.7)</td>
<td>100.2 (29.1)</td>
</tr>
<tr>
<td></td>
<td>81.1 (30.1)</td>
<td>79.6 (27.8)</td>
</tr>
<tr>
<td>BDI</td>
<td>20.1 (12.2)</td>
<td>19.4 (13.1)</td>
</tr>
<tr>
<td></td>
<td>17.3 (14.5)</td>
<td>20.0 (22.1)</td>
</tr>
<tr>
<td>BAI</td>
<td>17.6 (12.5)</td>
<td>18.6 (12.2)</td>
</tr>
<tr>
<td></td>
<td>16.1 (12.8)</td>
<td>17.0 (20.1)</td>
</tr>
</tbody>
</table>

Note: Statistical significance is denoted by $p$ values, with $p < 0.0001$ indicating highly significant differences.
Secondary Treatment Outcomes

Reduction in hoarding-associated functional impairment (total sample). We examined total scores on the ADL-H, which assesses functional impairment in daily living caused by hoarding behaviors, as a secondary outcome measure at baseline and posttreatment. Posttreatment data were available for 246 individuals. The mean pretreatment ADL-H score for the entire sample was 31.0 (N = 316; SD = 10.5). Participants had an 11.4% (4.4-point; SD = 7.8) improvement in ADL-H scores pretreatment to posttreatment, with a mean posttreatment ADL-H score of 25.8 (SD = 9.4).

CBT versus PFT. Baseline data for the ADL-H were not significantly different between CBT (mean scores = 29.9; SD = 9.5) and PFT (mean scores = 32.0; SD = 11.2) (Table 4). Participants in CBT had a 10.5% reduction in ADL-H scores (mean posttreatment score = 25.5; SD = 10.1), while participants in PFT had a 12.4% reduction in ADL-H scores (mean posttreatment score = 26.1; SD = 8.7) (Table 6). The test of noninferiority, assuming a between-group difference of ≤ 2.5, was significant (between-group difference = 1.82 points; 95% CI; t = –1.52; df = 1; p = 0.05), indicating equivalence between the groups. As with SI-R scores, the results did not change when analyses were limited to treatment completers only or when individuals who dropped out and restarted the group were excluded. For the ADL-H, the respective effect sizes were 0.60 (CBT) and 0.52 (PFT).

Predictors of treatment response. We examined predictors of treatment response for the posttreatment assessment only; we did not include the longitudinal assessment in this analysis. Baseline hoarding-related measures, particularly more severe pretreatment severity (t = 3.75; df = 1; p < .0001) and hoarding-related thoughts and beliefs (t = 2.50; df = 1; p = .013), were associated with treatment improvement. More severe anxiety and depression were associated with less improvement on the ADL-H but not on the SI-R. Greater attendance at group sessions and homework completion were associated with treatment improvement for both treatment groups, with a stronger effect in PFT than in CBT.

In multivariable models we included variables associated with change in either SI-R or ADL-H scores at a P value of ≤ 0.15 using linear regression for any of the univariable analyses
(ie, in the entire sample, in the G-CBT sample alone, or in the G-PFT sample alone). These included years of education ($p = 0.15$ for ADL-H), hoarding severity ($p < 0.0001$ for SI-R), number of psychiatric diagnoses in addition to HD ($p = 0.07$ for ADL-H), hoarding-related thoughts and beliefs (SCI; $p = 0.0005$ for ADL-H), depression (BDI; $p = 0.08$ for ADL-H) and anxiety (BAI; $p = 0.01$ for ADL-H) scores, attention and concentration (SNAP-IV; $p = 0.007$ for ADL-H), dementia symptoms (MoCA; $p = 0.05$), finding a clutter buddy helpful ($p = 0.01$ for SI-R), preferring a mental health provider to a peer facilitator ($p = 0.04$ for SI-R), percentage of homework completed ($p = 0.001$ for SI-R), and percentage of groups attended ($p = 0.003$ for SI-R). For highly correlated variables, we included only the most strongly associated in the univariate analyses, resulting in the removal of some that passed screening. We used the same variables in both models of SI-R and ADL-H change. Variables included in the final multivariable models were pretreatment UHSS, BDI, BAI, SNAP-IV, and SCI scores; homework adherence; and how helpful participants found their clutter buddy to be. The final multivariable model explained only a small proportion of the change in SI-R scores pretreatment to posttreatment (adjusted $R^2 = 0.072$) (Table 7). Only pretreatment hoarding severity and homework adherence were independently associated with improvement. We saw similar results for change in ADL-H scores; that is, the test of the overall model was significant, although no individual variables were independently associated with treatment improvement, and again, the model accounted for only a small proportion of the pretreatment to posttreatment change (N = 237; adjusted $R^2 = 0.036; f^2,229 = 2.38; p = 0.023$). These results did not change when we repeated the analyses in treatment completers only. Pretreatment hoarding severity was the only independent predictor of treatment outcome in CBT ($\beta = 0.787; SE = 0.283; t = 2.78; df = 1; p = 0.006$); higher pretreatment severity was associated with greater improvement. For PFT, higher pretreatment hoarding severity ($\beta = 0.728; SE = 0.281; t = 2.59; df = 1; p = 0.011$) and better homework adherence were significant predictors of improvement ($\beta = 12.56; SE = 5.25; t = 2.39; df = 1; p = 0.018$), and lower pretreatment depressive symptoms ($\beta = -0.277; SE = 0.143; t = -1.93; df = 1, p = 0.057$) indicated a trend for association with improvement.
Table 7. Multivariable model for pretreatment to posttreatment total score change (SI-R) entire sample.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>df</th>
<th>Parameter</th>
<th>t Statistic</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCI pretreatment score</td>
<td>1</td>
<td>−0.014</td>
<td>−0.34</td>
<td>0.732</td>
</tr>
<tr>
<td>UHSS pretreatment score</td>
<td>1</td>
<td>0.746 (0.199)</td>
<td>3.75</td>
<td>&lt;0.00001</td>
</tr>
<tr>
<td>BAI pretreatment score</td>
<td>1</td>
<td>0.087 (0.101)</td>
<td>0.86</td>
<td>0.391</td>
</tr>
<tr>
<td>BDI pretreatment score</td>
<td>1</td>
<td>−0.134</td>
<td>−1.33</td>
<td>0.184</td>
</tr>
<tr>
<td>SNAP-IV pretreatment</td>
<td>1</td>
<td>0.035 (0.072)</td>
<td>0.49</td>
<td>0.627</td>
</tr>
<tr>
<td>Homework adherence (%)</td>
<td>1</td>
<td>10.023 (4.00)</td>
<td>2.50</td>
<td>0.013</td>
</tr>
<tr>
<td>Found clutter buddy helpful</td>
<td>1</td>
<td>0.192 (0.766)</td>
<td>0.250</td>
<td>0.802</td>
</tr>
<tr>
<td>Intercept</td>
<td>1</td>
<td>−9.327</td>
<td>−1.51</td>
<td>0.133</td>
</tr>
</tbody>
</table>

Abbreviations: BAI, Beck Anxiety Inventory; BDI, Beck Depression Inventory; SCI, Saving Cognition Inventory; SNAP-IV, Swanson, Nolan and Pelham ADHD Rating Scale; UHSS, UCLA Hoarding Severity Scale.

Table Note: N = 238; F2740; adjusted R² = 0.072; p = 0.001. Found clutter buddy helpful scoring: 1 = Very helpful to 4 = Unhelpful. Those with no clutter buddies were coded as 4. Negative parameter estimates indicate that higher scores are associated with less improvement; positive estimates indicate that higher scores are associated with more improvement.

Longitudinal treatment outcomes. Although initiated in the second year of the study, all longitudinal follow-up assessments were conducted in the same manner as the original assessments and blinded to treatment group. We included only individuals who were randomized for treatment and provided permission for recontact. We obtained longitudinal follow-up data (defined as second posttreatment SI-R and ADL-H scores at least 3 months following completion of treatment) on 183 of the 323 randomized participants (101 in CBT and 82 in PFT) (X² = 5.40; df = 1; p = 0.02). The average time to follow-up was 14.4 months (range 3-25). Of participants, 36% were recontacted within a year of completing treatment, and all but one were recontacted within 2 years; one participant responded 25 months after completing treatment. There was no significant difference between mean time to follow-up between CBT (mean = 13.9 months; SD = 7.2) and PFT (mean = 13.2; SD = 6.9) participants (t = −0.67; df = 1; p = 0.50). Participants who provided longitudinal follow-up data were more likely to be women (X² = 9.3; df = 1; p = 0.002) and white (X² = 11.8; df = 1; p = 0.001) and had slightly higher education levels (t = −3.6; df = 1; p = 0.0004) than those who did not. There were no significant differences in pretreatment or posttreatment SI-R scores between those for whom we obtained longitudinal data and those for whom we could not obtain such data. Of the 183 individuals for whom we had information, 74.5% remained in contact with or continued to meet with the
members of their treatment group after completing the study, while 31% remained in contact with their clutter buddies. Twenty-three percent attended drop-in or other support groups for HD, 17.5% received ongoing or additional formal treatment for HD (eg, by a psychiatrist or psychologist), and 5% had worked with a professional organizer. Thirty-eight percent continued to receive help with hoarding symptoms from family and/or friends.

The mean SI-R score at longitudinal follow-up was 48.1 (SD = 14.4), and the mean ADL-H score was 27.7 (SD = 10.8) for the entire sample. There were no significant differences between SI-R scores immediately posttreatment and at longitudinal follow-up (mean SI-R difference = –0.45; SD = 12.7; t = –0.45; df = 1; p = 0.66), indicating that the treatment gains were generally maintained at 3 months post treatment or later (Figure 2). The linear mixed models analysis incorporating treatment group (CBT, PFT) and time (pretreatment, posttreatment, and longitudinal) was significant for both SI-R (LR X² = 115.5; df = 1; p < 0.0001) and ADL-H (LR X² = 162.1; df = 1; p < 0.0001). For both outcomes, only time significantly contributed to the model (SI-R: F = 250.5, df = 2, p < 0.0001; ADL-H: F = 34.3, df = 2, p < 0.0001). Neither group nor the group–time interaction variables were statistically significant.

There were no significant differences in longitudinal SI-R (t = –0.15; df = 1; p = 0.88) or ADL-H (t = 0.09; df = 1; p = 0.93) scores between the CBT and PFT groups (Table 6). Mean ADL-H scores in the entire sample were higher at longitudinal follow-up than immediately posttreatment (ADL-H difference = –2.2; SD = 10.5; t = –2.68; df = 1; p = 0.008), indicating that gains in function were not as well sustained as was hoarding symptom improvement. Longer time from posttreatment to longitudinal follow-up was associated with worsening SI-R scores (r = –0.173; p = 0.03) and ADL-H scores (r = –0.240; p = 0.01). Of those who continued to receive some type of help with hoarding behaviors posttreatment, only ongoing help from family and/or friends was significantly associated with maintenance of gains. Individuals who received such help had mean longitudinal SI-R scores of 43.9 (SD = 13.3), while those who did not had mean longitudinal SI-R scores of 50.2 (SD = 14.7) (t = 2.89; df = 1; p = 0.004; N = 177). The remission rate (defined by SI-R < 42 and SI-R delta ≥ 14) for the longitudinal follow-up was 30.1%; 32.8% of CBT and 29.4% of PFT participants (X² = 0.33; p = 0.56). There were no statistically significant differences between these groups for sex, age, pretreatment hoarding
severity or cognitions, psychiatric comorbidity, treatment completion, group attendance rates, or homework completion rates (data not shown). Twenty-eight participants who achieved remission at posttreatment no longer met criteria for remission at longitudinal follow-up, while 24 participants who did not meet criteria for remission at posttreatment did meet these criteria at longitudinal follow-up.

**Figure 2.** Hoarding symptom severity scores (SI-R total score) before (PRE-TX), immediately following (POST-TX), and at least 3 months after treatment (F/U). Change in SI-R total score for CBT compared with PFT.

**Patterns of treatment response:** We next classified respondents into 3 categories:

1. **respond:** individuals who showed at least clinically meaningful improvement with treatment (change in SI-R total score ≥ 10 points);
2. **no change:** those whose posttreatment scores were between 9 points below and 4 points above their pretreatment SI-R scores;
3. **worse:**
those whose symptoms worsened after treatment (conservatively defined as worsening in SI-R score of ≥ 5 points). Of the 247 individuals with posttreatment data, 70.4% responded to treatment (N = 173), 25.5% had no change (N = 63), and 4.5% got worse (N = 11). Of those with available longitudinal data, 81.7% (N = 94/115) of the treatment responders maintained their gains, while 32.5% (N = 13/40) of those with no change and 12.5% (N = 1/8) who got worse at posttreatment were treatment responders at the longitudinal follow-up ($X^2 = 44.8; df = 2; p < 0.0001$). Those who were responders at the posttreatment assessment had somewhat higher pretreatment SI-R scores ($F = 4.73; df = 2; p = 0.01$), but there were no other significant differences between the groups.

**Neuropsychological tests.** We used multiple regression to examine whether neurocognitive function predicted treatment outcome, as well as whether participation in PFT or CBT was associated with improvement in specific aspects of cognitive function. When we included all assessments in the model, neuropsychological scores did not explain a significant amount of the variance in posttreatment SI-R scores (N = 186; adjusted $R^2 = 0.0; F = 1.07; p = 0.383$). The only marginally significant independent predictor was the SDMT, which measures speed of information processing ($\beta = –0.945; SE = 0.485; t = –1.95; p = 0.053$). When we used the ADL-H as the dependent variable, the final model accounted for a nonsignificant proportion of the total variance (N = 185; adjusted $R^2 = 0.01; F = 1.13; p = 0.327$). Only the second part of the Delis–Kaplan Executive Function System’s sorting test, which measures executive functioning, was a significant predictor ($\beta = –0.838; SE = 0.327; t = –2.56; p = 0.011$). The results were similar when we examined CBT and PFT participants separately.

Next, we conducted t tests to examine whether participation in CBT or PFT was associated with improvement in specific cognitive domains. Both CBT (Table 8) and PFT (Table 9) participation was associated with significant improvement across several areas of cognitive functioning.
Table 8. Neurocognitive Mean Scaled Scores Pretreatment and Posttreatment, Mean Difference, and \( t \) Test \( P \) Values for CBT Treatment Group

<table>
<thead>
<tr>
<th>NP Tests</th>
<th>Pre – Mean SS</th>
<th>Post – Mean SS</th>
<th>Mean Change SS</th>
<th>( t ) Test ( P ) Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>HVLTL Learning Trials</td>
<td>8.9</td>
<td>9.0</td>
<td>−0.024</td>
<td>0.926</td>
</tr>
<tr>
<td>HVLTL Delayed Recall</td>
<td>9.0</td>
<td>9.3</td>
<td>−0.041</td>
<td>0.874</td>
</tr>
<tr>
<td>HVLTL Retention %</td>
<td>9.6</td>
<td>10.2</td>
<td>−0.252</td>
<td>0.429</td>
</tr>
<tr>
<td>HVLTL Recognition</td>
<td>9.5</td>
<td>9.2</td>
<td>0.246</td>
<td>0.471</td>
</tr>
<tr>
<td>Discrimination Index</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CPT Perseverations</td>
<td>8.0</td>
<td>8.6</td>
<td>−0.177</td>
<td>0.658</td>
</tr>
<tr>
<td>SDMT</td>
<td>8.4</td>
<td>9.1</td>
<td>−0.605</td>
<td><strong>0.0013</strong></td>
</tr>
<tr>
<td>Stroop Word</td>
<td>7.9</td>
<td>8.0</td>
<td>0.238</td>
<td>0.221</td>
</tr>
<tr>
<td>Stroop Color</td>
<td>7.2</td>
<td>7.6</td>
<td>−0.115</td>
<td>0.511</td>
</tr>
<tr>
<td>Stroop Color-Word</td>
<td>9.2</td>
<td>10.1</td>
<td>−0.754</td>
<td><strong>&lt;.0001</strong></td>
</tr>
<tr>
<td>Stroop Interference</td>
<td>9.5</td>
<td>10.3</td>
<td>−0.771</td>
<td><strong>0.0001</strong></td>
</tr>
<tr>
<td>BVMT-R Total Recall</td>
<td>8.1</td>
<td>10.0</td>
<td>−1.569</td>
<td><strong>&lt;.0001</strong></td>
</tr>
<tr>
<td>BVMT-R Learning</td>
<td>11.2</td>
<td>10.5</td>
<td>0.431</td>
<td>0.277</td>
</tr>
<tr>
<td>BVMT-R Delayed Recall</td>
<td>8.9</td>
<td>10.1</td>
<td>−0.878</td>
<td><strong>0.012</strong></td>
</tr>
<tr>
<td>Block Design</td>
<td>10.2</td>
<td>11.0</td>
<td>−0.691</td>
<td><strong>&lt;.0001</strong></td>
</tr>
<tr>
<td>Matrix Reasoning</td>
<td>11.3</td>
<td>12.0</td>
<td>−0.469</td>
<td><strong>0.064</strong></td>
</tr>
<tr>
<td>Digit Span Forward</td>
<td>10.3</td>
<td>10.7</td>
<td>−0.146</td>
<td>0.492</td>
</tr>
<tr>
<td>Digit Span Backwards</td>
<td>10.4</td>
<td>11.2</td>
<td>−0.438</td>
<td><strong>0.029</strong></td>
</tr>
<tr>
<td>Sorting Card Set 1/3</td>
<td>10.6</td>
<td>11.1</td>
<td>−0.164</td>
<td>0.512</td>
</tr>
<tr>
<td>Sorting Card Set 2/4</td>
<td>10.6</td>
<td>9.7</td>
<td>0.967</td>
<td><strong>0.0002</strong></td>
</tr>
<tr>
<td>IGT Net Total</td>
<td>10.3</td>
<td>11.3</td>
<td>−0.857</td>
<td><strong>0.019</strong></td>
</tr>
</tbody>
</table>

Abbreviations: BVMT-R, Brief Visuospatial Memory Test, Revised; CPT, Connors Continuous Performance Test II; HVLTL, Hopkins Verbal Learning Test, Revised; IGT, Iowa Gambling Task; SDMT, Symbol Digit Modalities Test.
Table 9. Neurocognitive Mean Scaled Scores Pretreatment and Posttreatment, Mean Difference, and t Test P Values for PFT Treatment Group

<table>
<thead>
<tr>
<th>NP Tests</th>
<th>Pre – Mean</th>
<th>Post – Mean</th>
<th>Mean Change</th>
<th>t Test P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>HVLT Learning Trials</td>
<td>9.0</td>
<td>9.6</td>
<td>−0.246</td>
<td>0.282</td>
</tr>
<tr>
<td>HVLT Delayed Recall</td>
<td>9.2</td>
<td>9.9</td>
<td>−0.415</td>
<td>0.102</td>
</tr>
<tr>
<td>HVLT Retention %</td>
<td>9.8</td>
<td>10.4</td>
<td>−0.686</td>
<td>0.035</td>
</tr>
<tr>
<td>HVLT Recognition Discrimination Index</td>
<td>9.3</td>
<td>9.5</td>
<td>0.153</td>
<td>0.580</td>
</tr>
<tr>
<td>CPT Perseverations</td>
<td>8.0</td>
<td>8.9</td>
<td>−0.620</td>
<td>0.147</td>
</tr>
<tr>
<td>SDMT</td>
<td>8.8</td>
<td>9.5</td>
<td>−0.470</td>
<td>0.028</td>
</tr>
<tr>
<td>Stroop Word</td>
<td>7.3</td>
<td>7.2</td>
<td>−0.050</td>
<td>0.765</td>
</tr>
<tr>
<td>Stroop Color</td>
<td>7.0</td>
<td>7.4</td>
<td>−0.350</td>
<td>0.024</td>
</tr>
<tr>
<td>Stroop Color-Word</td>
<td>9.4</td>
<td>10.0</td>
<td>−0.432</td>
<td>0.011</td>
</tr>
<tr>
<td>Stroop Interference</td>
<td>10.0</td>
<td>10.4</td>
<td>−0.205</td>
<td>0.190</td>
</tr>
<tr>
<td>BVMT-R Total Recall</td>
<td>8.7</td>
<td>10.5</td>
<td>−1.513</td>
<td>&lt; .0001</td>
</tr>
<tr>
<td>BVMT-R Learning</td>
<td>11.6</td>
<td>10.4</td>
<td>1.454</td>
<td>0.0002</td>
</tr>
<tr>
<td>BVMT-R Delayed Recall</td>
<td>9.5</td>
<td>10.7</td>
<td>−0.824</td>
<td>0.0016</td>
</tr>
<tr>
<td>Block Design</td>
<td>10.5</td>
<td>11.3</td>
<td>−0.874</td>
<td>&lt; .0001</td>
</tr>
<tr>
<td>Matrix Reasoning</td>
<td>11.6</td>
<td>12.3</td>
<td>−0.327</td>
<td>0.120</td>
</tr>
<tr>
<td>Digit Span Forward</td>
<td>10.3</td>
<td>10.5</td>
<td>−0.183</td>
<td>0.502</td>
</tr>
<tr>
<td>Digit Span Backwards</td>
<td>10.6</td>
<td>11.0</td>
<td>−0.441</td>
<td>0.062</td>
</tr>
<tr>
<td>Sorting Card Set 1/3</td>
<td>10.7</td>
<td>11.0</td>
<td>−0.240</td>
<td>0.394</td>
</tr>
<tr>
<td>Sorting Card Set 2/4</td>
<td>10.7</td>
<td>9.8</td>
<td>1.158</td>
<td>0.0002</td>
</tr>
<tr>
<td>IGT Net Total</td>
<td>10.1</td>
<td>10.3</td>
<td>−0.275</td>
<td>0.482</td>
</tr>
</tbody>
</table>

Abbreviations: BVMT-R, Brief Visuospatial Memory Test, Revised; CPT, Connors Continuous Performance Test II; HVLT, Hopkins Verbal Learning Test, Revised; IGT, Iowa Gambling Task; SDMT = Symbol Digit Modalities Test

*Tables required by and found on ClinicalTrials.gov and are provided at the end of the report.
Discussion

Decisional Context

In this randomized clinical trial of 323 participants with HD, peer-facilitated group treatment was as effective as psychologist-led group treatment in reducing hoarding-related symptoms, with an overall symptom reduction rate of 27% across both treatment groups. Treatment gains were maintained for at least 3 months. Independent predictors of treatment outcome included pretreatment hoarding severity and homework adherence. Continuing to receive support from family and friends was the only predictor of maintenance of gains at 3 months or later. These results have important implications for treatment providers, for families and individuals with HD, and for policymakers. Although effective, the availability of trained mental health professionals who can provide treatment for HD is limited, and access to care remains a substantial impediment for many individuals and families.

The finding that community-based peer-facilitated group treatment was as effective as treatment with trained mental health professionals represents a potential paradigm shift in treatment for HD, and has the potential to substantially increase access to care for individuals living with HD. Mental health organizations, academic and community treatment centers, and other relevant stakeholders can use the information obtained from this study to learn more about HD and effective ways to treat it, and possibly offer the peer-led treatment groups. Based on the study’s findings, researchers can develop more precise research questions in order to improve treatment and the overall understanding of HD.

The Study Results in Context

The results of this study confirm and expand the evidence base showing that CBT conducted by trained mental health professionals (in this case, psychologists) was effective in treating HD in a time-limited format, and for the first time, clearly demonstrates that peer-led treatment was equally effective. The prior published evidence suggests this outcome, but does not directly examine it in a randomized clinical trial. Our results also confirm and refine the average rates of improvement that can be expected from these forms of treatment (eg, 25% to 30% symptom improvement). Finally, we document, in a large sample, rates of response
(~66%) and remission (~30%), paralleling what has been found previously\textsuperscript{48}. The only previously published study to directly compare these 2 forms of treatment, and the catalyst for the current study, was conducted by our group retrospectively on data collected for clinical rather than research purposes, and therefore participants were not randomized and participants in each of the group treatments had clear differences in baseline characteristics.\textsuperscript{49}

The 27% average overall improvement in symptom severity found in this study across both types of treatment is in line with the 20% to 30% improvement in severity reported in previous trials testing CBT approaches.\textsuperscript{48-50} Two-thirds of participants had a clinically significant treatment response (≥14-point improvement), suggesting that most individuals benefited from treatment, although there is still a long way to go, as only one-third of participants achieved symptomatic remission as defined here.

We also examined independent and group predictors of response to treatment. We hypothesized that individual characteristics, including psychiatric comorbidity; insurance status; preferences for treatment type; and measures of visual categorization, learning, and memory would be associated with treatment outcome and could be used in the future to identify individuals who may be more or less likely to respond to a particular type of treatment. Contrary to our hypothesis, however, the only participant-related variable that consistently predicted treatment response in the entire sample was more severe hoarding symptoms at baseline, while treatment-related variables, in particular ongoing participation in hoarding-related services (including involvement of family and friends, participation in support groups, or formal treatment) were associated with long-term outcome. Psychiatric symptoms, neurocognitive status, and individual demographic characteristics (eg, sex, socioeconomic status, insurance status) were not predictive of treatment outcome, nor were participant beliefs about treatment efficacy or treatment preferences. Similarly, no participant-related variables independently predicted differential responses to CBT or PFT. However, treatment adherence (defined as either group attendance or homework adherence) did independently predict treatment outcome in the multivariable models. These findings reinforce the importance of identifying ways to improve treatment adherence and promoting ongoing community engagement (both during and after treatment) in order to maximize effectiveness.
and promote ongoing maintenance of gains for this chronic disorder. They also suggest that both CBT and PFT are robust to variations in administration and adherence, as despite several differences in protocol, there were no significant differences in outcomes between the 2 groups. Our findings are also in line with previous studies, which have not to date identified any individual characteristics that consistently are associated with treatment outcome. Our results underscore the importance of the ongoing involvement of family members and/or friends in the recovery process. Finally, we note that although there were no differences in the rates of treatment completers between groups, retention in the study for posttreatment assessments was consistently higher for the CBT than for the PFT groups, as was adherence (attendance and homework completion). It is currently unclear whether this finding is study specific, or whether it represents a higher degree of engagement in the CBT group in general, which, if replicated, would have important implications for dissemination of the treatment approaches.

While neurocognitive function at baseline did not predict treatment outcome, some of the neurocognitive measures did show evidence of improvement following treatment (Tables 8 and 9). Participants in both CBT and PFT showed significant improvement in speed of information processing, visual memory, and visuospatial processing. Tests measuring visual categorization showed worsening after treatment, somewhat surprisingly. Although these findings require further investigation, early analyses indicate that the CBT and PFT treatment groups differed in how treatment impacted other areas of cognitive functioning. Participants from the CBT groups showed significant improvement in measures of response inhibition, working memory, and decision making, whereas participants in the PFT groups did not. Participants in the PFT groups had significant improvement in verbal learning and memory, visual learning, and working memory. We note that these results are exploratory, and at this point suggest areas for further research rather than firm conclusions.

**Implementation of Study Results**

Both CBT and PFT treatment formats could easily be implemented into a typical care setting. HD-specific CBT led by licensed mental health providers is currently standard of care, and training of psychologists and other mental health providers in this modality can be
implemented into standard evidence-based psychotherapy trainings. Peer-facilitated group treatments, such as the PFT approach studied here, are already being offered by our community partner, MHASF, and other advocacy groups in other cities, and can be implemented in other community or treatment settings in a similar fashion. Further, working with stakeholders (MHASF) contributed greatly to the study in that the collaboration helped secure participant recruitment, treatment sites, and exchange of perspectives, which positively impacted the study’s protocol. Other mental health organizations and researchers could potentially use our study as a model of a successful partnership for research and/or clinical care.

**Generalizability**

We specifically designed this study to be generalizable to the larger community of individuals with HD. We limited our exclusion criteria to those that would impede participation in group treatment and included individuals with a wide range of psychiatric disorders and other potential challenges (including mild cognitive impairment). We also made substantial efforts to include individuals with a diversity of backgrounds and characteristics, including, but not limited to, sex, age, race, insurance status, socioeconomic status, and education. We were fairly successful in this regard—for most variables (with the notable exception of sex), our participants were representative of the population of the larger San Francisco Bay Area, which is itself a very diverse population. Therefore, our findings are likely to be broadly generalizable within the larger community of individuals with HD throughout the United States. An exception to this is that this study may not generalize to individuals with poor or absent insight into their HD, as we were unable to recruit into the treatment study individuals with absent insight and recruited relatively few with poor insight.

**Subpopulation Considerations**

We did not conduct separate analyses for specific subgroups, although as a part of our primary analyses we examined, in univariate analyses, the relationships between individual participant characteristics (including those defined by sex, race, ethnicity, age, psychiatric status, and socioeconomic status) and treatment outcomes. We did not find any differences in treatment outcomes for any of these variables. As noted above, this suggests that our findings
are widely generalizable to the broad community of individuals with HD in the United States.

**Study Limitations**

We designed this study to be implemented in a real-world setting, and as such, the study has multiple potential limitations. Although we had a high retention rate in general, about 30% of participants dropped out of treatment prior to completion, and we were able to obtain posttreatment data on only 16 of the participants who dropped out of treatment. Similarly, the longitudinal assessment, including the assessment of specific groups categorized by type of response, was designed and executed in the final year of the study rather than being planned prior to study initiation. As a result, participants were assessed at varying time points following treatment, and we obtained longitudinal follow-up data on only 55% of all randomized participants; 84% of the treatment completers provided longitudinal data. Finally, it is notable that there were substantial differences between treatment groups in those who provided the requested longitudinal follow-up data (N = 82 for PFT and N = 101 for CBT), which may speak to increased engagement among the CBT participants, as noted above. The results of this study should be interpreted keeping these limitations in mind.

As with many studies, since participants typically sought out the hoarding treatment and volunteered their time, participants in our study may have had a higher awareness of and knowledge base regarding their disorder than those who did not seek out treatment. As suggested by the data in Table 3, those who did not seek treatment may have had more mild HD or less insight into their illness, and therefore did not see it as problematic, as has been posited in other HD-related research. Although we made every effort to make the treatment facilities accessible to those with limited mobility, we were unable to reach and/or treat individuals who were confined to their homes or for other reasons could not attend treatment groups. Despite our diverse recruitment approaches, most of the participants in the study were female, as has been seen in previous studies, and most were white, although the ethnic and racial makeup of the sample paralleled that of the San Francisco Bay Area as a whole. Thus, future studies should focus on underrepresented groups to ensure generalizability of the findings.

Finally, there were several potential limitations to the study design. First, the outcome
variables were dependent on self-report measures rather than objective observations. Second, we made several modifications to the published manuals to fit the needs of our intended population and treatment providers. We shortened the recommended length of the CBT treatment to better match the PFT groups, we did not formally assess adherence to the treatment manual, and we did not conduct home visits prior to study entry to assess the degree of clutter present; rather, we relied on self-report. These changes were made to facilitate the treatment delivery to as many individuals as possible in a real-world setting. For example, because home visits are not part of standard clinical care for many mental health providers, this is consistent with our real-world approach. However, it is possible that some participants were inaccurate in their self-reports of clutter. Thus, these changes represent both a potential weakness of the study, but also a potential strength, as we demonstrated that a substantial proportion of our participants improved under the protocol described here and that the 2 treatment types did not significantly differ in outcomes.

**Future Research**

Future research efforts should focus on improving treatment outcomes in both treatment arms. Although we had good response rates, we achieved functional remission in only 30% of participants, and the average improvement was 27% in severity ratings. Similarly, we had an only 11% improvement in hoarding-related impairment as measured by the ADL-H; this is approximately half of what has been seen in previous studies, despite reductions in self-reported symptom severity that are similar to previous reports. Additional work is clearly needed to refine the treatment protocols, to examine additional predictors of treatment response, and to identify appropriate and effective adjunct treatment options. Additional research should also focus on dissemination of the peer-facilitation model, including the role and appropriate intensity/frequency of training and ongoing supervision/consultation for peer facilitators, utility in structured settings such as senior centers or assisted living facilities, and utility for those with limited or poor insight into their illness.
Conclusion

Workbook-based group treatment led by peer facilitators was as effective in treating hoarding disorder as psychologist-led treatment groups. This finding provides substantial evidence for an effective, alternative treatment avenue for individuals who do not have access to care by mental health professionals or who prefer not to seek care in a mental health setting. Having a more accessible treatment available for individuals suffering from HD can help a greater number of people benefit from treatment. By doing this, it can also create a larger community for individuals suffering from HD and offer a platform for connecting with and receiving support from others suffering from the same disorder. Peer-facilitated treatment has the potential to reduce stigma around HD and create more a supportive environment for those suffering from this chronic and impairing disorder.

Another important finding was the relationship between treatment adherence (defined either as group attendance or homework adherence), family support, and improved treatment outcomes, both short and long term. This highlights the importance of developing innovative methods to improve adherence among patients and families.

Further research could examine the effectiveness of peer-led groups for individuals with other disorders, further review strategies for patient treatment adherence, and work with a community support organization.
References


Publications

Published Journal Articles:


Appendix

Appendix 1

CBT Session Topics

Session 1: Introduction and psychoeducation
Session 2: Cognition, emotional attachment to saved items
Session 3: Motivation
Session 4: Exposure Model
Session 5: Practice, home commitment
Session 6: Organization, practicing making decisions
Session 7: Barriers to making progress
Session 8: Paper filing organization
Session 9: Letting go: Direct exposure
Session 10: Family and letting go: Instructions for caches or “clutter buddies”
Session 11: Maintenance and Letting Go
Session 12: Acquisition
Session 13: Acquisition continued
Session 14: Maintenance and relapse prevention
Session 15: Revisiting key topics
Session 16: Termination and next steps

BiT Session topics

Session 1: Getting to know each other
Session 2: Meet the Bad Guys
Session 3: Here Come the Good Guys!
Session 4: How Did This Happen?
Session 5: Enhancing Motivation
Session 6: Help With Reducing Acquiring
Session 7: More Help With Reducing Acquiring
Session 8: Sorting and Discarding: Getting Ready
Session 9: Sorting and Discarding: Let’s Go!
Session 10: Working with clutter buddies
Session 11: Sorting and Discarding: Practicing Letting-Go
Session 12: Sorting and Discarding: Succeeding!
Session 13: Here Come the Bad Guys again
Session 14: Taking On Your Brain
Session 15: Maintaining Success
Appendix 2

Beliefs About Group Treatment for Hoarding and Cluttering

This questionnaire asks about your thoughts on what aspects of group therapy for hoarding and cluttering you think will be the most helpful to you. There are no right or wrong answers, just answer as honestly as you can. All answers will be kept private and confidential.

For the following statements, circle the number that best corresponds with how you feel.

<table>
<thead>
<tr>
<th>Strongly agree</th>
<th>Somewhat agree</th>
<th>Neither agree</th>
<th>Somewhat disagree</th>
<th>Strongly disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td>5--------------</td>
<td>4--------------</td>
<td>3--------------</td>
<td>2-----------------</td>
<td>1-----------------</td>
</tr>
</tbody>
</table>

1. Participating in a group treatment will work as well as individual treatment for me.

2. I work best with a licensed mental health provider.

3. I work best with others who have personal experience with hoarding and cluttering challenges.

4. I believe that cognitive-behavioral therapy, which follows a structured plan and asks me to complete homework assignments, will best fit my style.

5. I believe that a peer-lead support group that follows the readings and assignments from a book will best fit my style.

6. If given the choice, I would prefer participating in a cognitive-behavioral treatment group provided by a mental health professional for treatment of my hoarding and cluttering challenges.

7. If given the choice, I would prefer participating in a peer-facilitated support group treatment of hoarding and cluttering challenges.
Appendix 3

Weekly Group Evaluation
Your responses to these questions are confidential

1. How would you evaluate the quality of the treatment group?
   ( ) Excellent ( ) Good ( ) Fair ( ) Poor

2. Was the group beneficial to you?
   ( ) Yes ( ) No ( ) Not Sure
   If you answered no or not sure, please explain briefly on the back of this form

3. How would you rate the severity of your hoarding/cluttering problems now?
   ( ) Mild ( ) Neither Severe nor Mild ( ) Severe ( ) Very Severe

4. How does this compare to when you first started the group?
   ( ) Better ( ) The Same ( ) Worse ( ) Not Sure ( ) First Time

5. Thinking back to the last time you attended the group, have you done anything to reduce your hoarding/cluttering?
   ( ) Yes ( ) No ( ) Not Sure ( ) First Time
   If you answered no or not sure, please explain briefly on the back of this form

6. Thinking back to when you first began to attend the group, have you reduced the amount of clutter in your house?
   ( ) Yes, Very Much ( ) Yes, Somewhat ( ) No ( ) Not Sure ( ) First Time

7. Have you met with or spoken to your clutter buddy this week?
   ( ) Yes ( ) No ( ) Does not apply

8. How helpful was your clutter buddy in assisting or motivating you to complete your group assignment or reduce your clutter?
   ( ) Very Helpful ( ) Somewhat helpful ( ) Not very helpful ( ) Unhelpful ( ) Does not apply

9. How do you feel today about your overall prospects for recovery from your hoarding and cluttering challenges compared to before you started the group?
   ( ) Very hopeful ( ) Somewhat hopeful ( ) No change ( ) Somewhat less hopeful ( ) Much less hopeful

10. How would you rate the hoarding and cluttering treatment group overall?
    ( ) Excellent ( ) Good ( ) Fair ( ) Poor

11. Have you started any new psychiatric medications over the last two weeks? ( ) Yes ( ) No
    • What kind/type?

12. Have you started any new mental health treatment over the last two weeks? ( ) Yes ( ) No
    • What kind/type?

Do you have any comments or suggestions regarding the group? Please use the back of this form.
Appendix 4
Continuing Support Questionnaire

Since you have finished the group, have you:

Stayed in contact with your clutter buddy (if you had one)? Y/N
  Are you still in contact with your clutter buddy? Y/N

Stayed in contact with group members and/or continued to meet with your group? Y/N
  Are you still in contact with group members and/or continuing to meet with your group? Y/N
Attended drop in groups or any other support groups for hoarding? Y/N ______________________
  What drop in group or other support groups have you attended? ______________________
  Are you still attending the drop in or other support groups for hoarding? Y/N

Received any additional formal treatment for hoarding? For example, received treatment from a
psychiatrist or psychologist? Y/N
  If so, what kind of treatment was it? ______________________
  Are you still receiving that treatment? Y/N

Worked with a professional organizer? Y/N
  Are you still working with a professional organizer? Y/N

Had friends or family members (not your clutter buddy) help you to declutter? Y/N

Note: questions are asked in a branching fashion, such that if the first question is answered
“yes”, the following questions (indented) appear.