Cognitive Rehabilitation and Exposure Therapy for Veterans with Hoarding Disorder

NCT02402647

February 1, 2023

Current Proposal

Proposed Study Design

The proposed study includes a 26-session controlled, randomized clinical trial designed to test the hypothesis that while targeting avoidance produces changes in HD symptoms, targeting executive dysfunction as well as avoidance will lead to even greater symptomatic changes, and secondarily in functioning and quality of life. Further, we will explore outcome moderators (age and baseline executive functioning) and time to maximum effect. To do this, we will compare CREST (CCT + ET) to ET alone. The study will be comprised of 4 phases (see Figure 2) (1) Pre-Treatment Evaluation Phase (1 assessment); (2) During Treatment Evaluation Phase (3 assessments; at sessions 7, 13, and 21); (3) End of Treatment (1 assessment after session 26); and 4) 3 and 6-month Follow-up Phase (2 assessments).

Patient Inclusion/Exclusion

The inclusion criteria are designed to enroll a sample of Veteran HD patients with some level of executive functioning impairment in order to engage treatment targets. We do not believe this will exclude a significant number of Veterans. In our most recent sample of 55 HD patients over age 60, 73% were impaired on one or more of four DKEFS subtests (Tower, Trail Making, Verbal Fluency, or Color-Word Interference Tests). For the proposed study, patients must demonstrate impaired performance on one or more of the five DKEFS subtests given (Color-Word Interference, Verbal Fluency, Trail Making, Design Fluency, Tower Test). These criteria are optimal in that it will allow for a range of executive functioning within our sample to look at predictors of response as well as allow us to engage this intermediary target.

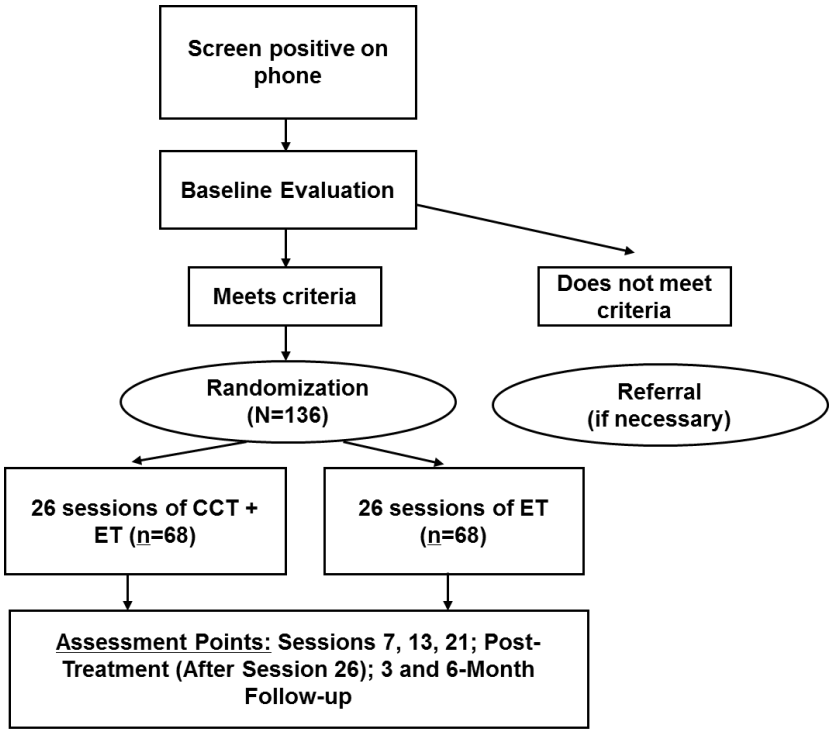


Figure 2. Study Design

Further, in order to ensure that we have an enriched ED sample, we will include individuals between the ages of 45-85. This older sample will allow for increase efficiency in finding Veteran participants with a level of ED, given that we know older adults with HD show ED[22,31], and that HD symptoms (which are correlated with ED) may become worse with age[45]. Given the mixed findings on ED in younger HD populations, this will prevent us from having a floor effect, thus ensuring that we have a sample where we can engage targets. Further, with this age range we will be able to examine how aging contributes to treatment outcomes, a gap in the existing research.

Other **inclusion criteria** include (1) Veterans age 45-85; (2) Hoarding Disorder diagnosis outlined by the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5)[6] as measured by the Structured Interview for Hoarding Disorder (SIHD)[67]; (3) HD as a primary diagnosis; (4) executive dysfunction as defined by an impaired scaled score (i.e., <7) on any of the following tests; DKEFS Color-Word Interference, Verbal Fluency, Trail Making, Design Fluency, Tower Test (5) stable on medications for at least 12 weeks, with no pharmacologic changes expected or made during the 12-month study (6) voluntary consent to participate. **Exclusion criteria** includes (1) diagnosis of a psychotic disorder, bipolar I, bipolar II disorder, substance abuse disorder as measured by the Mini-International Neuropsychiatric Interview (M.I.N.I.)[68] (3) current or history of any neurodegenerative disease; (4) active suicidal ideation; (5) concurrent participation in any form of psychotherapy, or prior ET, CBT for HD, or CREST for HD.

Recruitment and Randomization

Recruitment will take place in through the San Diego VA utilizing existing recruitment method for our HD studies including review of consults submitted to the VA San Diego Anxiety Disorders clinic, posted flyers in the hospital, VA clinician outreach, participation in the San Diego Hoarding Task Force (PI is a founding

member), and electronic media advertisements. The PI on the proposed project is the Program Director of the VA San Diego Anxiety Disorders Clinic, where Veterans with known HD symptoms are referred.

Our group has a successful history of recruiting HD patients. Since 2009, we have received over 270 geriatric (over age 60) HD referrals and since 2011 we have received over 125 midlife (under age 60) HD referrals. We have received an increase in calls (7 calls per week) since July 2013 due to our established reputation within the VA Healthcare System and San Diego County. Interestingly, through task force participation and electronic advertisements, we have been able to reach Veterans who were previously not receiving services at the VA and were able to enroll them in our local VA healthcare system. Our funded HD studies within the lab have been completed ahead of schedule due to availability of participants. Thus, we will be able to meet our enrollment goals with our existing recruitment system.

Recruitment targets will be 11-12 patients per study wave (see timeline below) for a total of 136 patients, half of whom will be randomly assigned to CREST. Participants will be administered the Hoarding Rating Scale (HRS)[69] over the phone by the study coordinator. If they score above a 4 on any item on the HRS, they will be scheduled for an in-person baseline assessment in their home to establish eligibility.

Following informed consent, an independent rater will complete a baseline assessment to determine eligibility. After review of the baseline scores, Drs. Ayers and Twamley (Co-I) will determine eligibility based on consensus. Dr. Golshan, the study statistician, will prepare the randomization table prior to recruitment of subjects. Subjects will be randomized 1:1 ratio to one of two conditions.

Retention

Our group currently has a markedly low attrition rate for HD studies. We have had 0% attrition across 2 open trials[15,33] and 8 in our current randomized controlled trial (3 for CREST; 5 for control). Participants will be compensated for their baseline, post-treatment and follow-up assessment points ($80 total; $20 per assessment) and a bonus of $20 for completing all assessments as an additional incentive to participate in treatment. We will optimize follow-up with regular patient contact in treatment, mailings to confirm whereabouts (e.g., birthday and holiday cards), and consent to contact at least one person likely to know subject whereabouts.

Treatment Conditions

Format and Course of Treatment Given the chronic and severe nature of HD, researchers have found that it requires a relatively lengthy course of treatment[14] (24-26 individual sessions)[57,70]. Both the control and experimental condition utilizes a 60 minute session-by-session manualized treatment protocol and will take 6 months to complete. Each participant will receive 26 sessions of individual treatment. In the experimental condition, patients will receive 7 sessions of CCT and 19 sessions of ET (consisting of exposure to discarding and acquiring modules). In the control condition, patients will receive 26 sessions of ET. The patient and therapist have their own workbook copy for both conditions to follow along during each session (Appendix 1 for CREST and Appendix 2 for ET condition). For the conceptualization of CREST, refer to Figure 3.

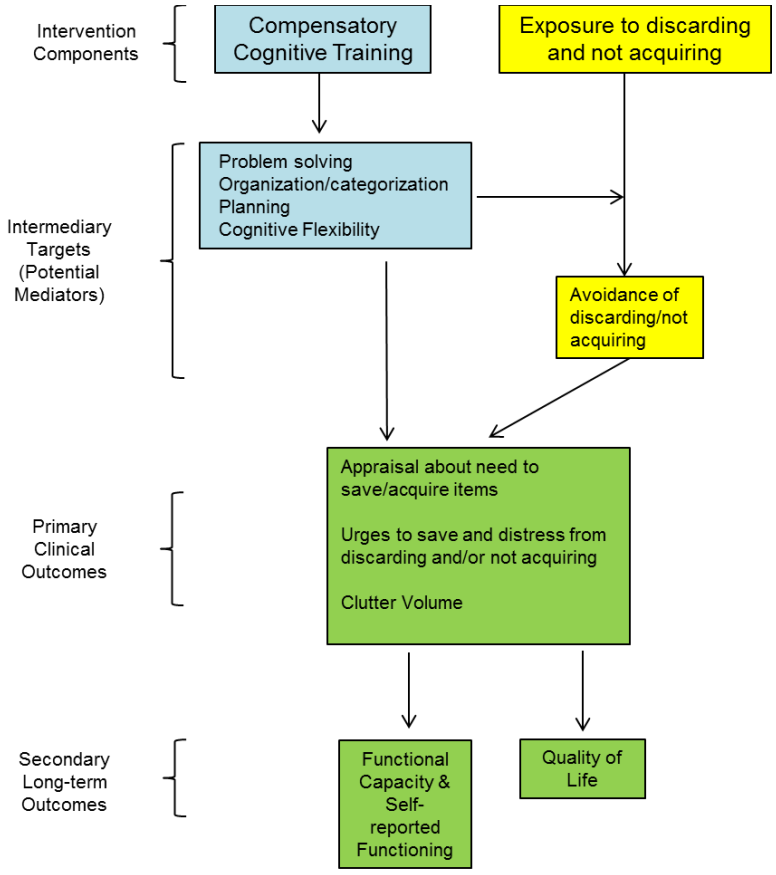


Figure 3. Conceptualization of CREST

Experimental Condition: CREST

Compensatory Cognitive Training (CCT) Modules (7 sessions). Compensatory Cognitive Training[60,61,62] is a manualized, low-tech, cognitive training intervention designed to target cognitive impairments common in people with psychiatric illness. The original CCT treatment manual has been distributed to over 250 clinicians and has been used with schizophrenia, brain injury, mild cognitive impairment, multiple sclerosis, Parkinson's

disease, and substance use disorder patients. The CCT modules specifically selected for CREST map onto known areas of HD neurocognitive deficits or weakness and include training in prospective memory, prioritizing, problem solving, planning, and cognitive flexibility. Because habit learning is also highly resistant to forgetting[71], we aimed to help Veteran participants form new habits in executive functioning to automate tasks and reduce the active cognitive effort usually demanded for effective performance. CCT does not use computers, and strategies taught do not "train to the test" or use any of the outcome measures during training. It is believed that using CCT will aide in the implementation of ET exercises. That is, CCT will provide patients with the necessary skills to fully comply with treatment (attendance and homework completion) and tools to prevent relapse.

A full outline of CCT modules is seen in Table 1. Given that HD patients are often disorganized and cannot manage their day to day demands, patients will be taught how to use a calendar system and practice utilizing their new scheduling skills. Specific details about calendar use are reviewed (e.g., when to check it, when to have a planning session, when to record dates, what to record, where it is kept, etc.).

Next, Veteran participants are taught how to start new behavioral routines by linking tasks (e.g., doing therapy homework after taking medications) and how to use a to-do list. Often HD clients are unrealistic about how many tasks they can accomplish in one day and the amount of time required for each item. To-do item creation, checking of the list, and scheduling to-dos are discussed and added to the calendar system.

Table 1. CREST Modules, Target Domains, and Manual Content

CCT Module (Target Domain)	**Session/Homework Content**
Session 1. Introduction to CCT and psychoeducation about the link between brain functioning and HD	Discussion of patient's experienced consequences of HD, barriers to treatment adherence and possible solutions, and treatment goals and expectations
Session 2. Calendar use (Prospective Memory)	Discussion of current/past calendar system; goal setting of where to keep calendar and when to use it; practicing use of calendar with "real-world" type scenario; using linking tasks and automatic places to help with daily activities
Session 3. Linking tasks, using a "to do" list (Prospective Memory)	Review of calendar use and linking tasks; using "to-do" lists along with calendar; determination of to do categories and frequencies; short-term prospective memory strategies (e.g., "can't miss reminders")
Session 4-5. Problem solving	Brainstorming practice exercises (therapist- and patient- provided examples); 6-step problem solving method; practice evaluating solutions for feasibility
Session 6. Thinking flexibly and planning (Cognitive Flexibility)	Self-talk and self-monitoring, brainstorming steps to meet a goal; practicing setting time lines for long-term goals.
Session 7. Home visit and organizational preparation (Cognitive Flexibility)	Creating a plan for organizing home (i.e., where types of items should be stored); brain storm "to-do" list for organization; rules of organizing; home maintenance system (current and proposed rules)
ET Module	**Session/Homework Content**
Session 8. Exposure preparation	Discussion of expectations, decision making, habituation, and avoidance; rules of discarding; motivational interviewing about willingness to change; acquisition hierarchy; fear of discarding hierarchy
Session 9. Introduction to exposure therapy	Review of organizational plan, rules of organizing, maintenance system, and rules of discarding; introduction to Subjective Units of Distress (SUDS) Ratings; discarding exposure
Sessions 10-22: Exposure to discarding & acquiring	Evaluation of progress; discarding exposure; discussion of reasons for saving and strategies for making discarding choices; introduction to advanced exposure (e.g., discussion of rules, preparing for sorting for a longer time and use of outside help); review of treatment goals
Sessions 23-24: Advanced exposure	Longer exposure time (2-4 hours) with outside help; Evaluation of progress; discussion of reasons for saving and strategies for making discarding choices; review of treatment goals
Sessions 25-26: Relapse prevention and maintenance	Review of progress in therapy; discussion of everyday uses for cognitive strategies (e.g., calendar use, "to do" lists, problem solving); exposure review; planning for the future

Problem solving skills are then taught as HD patients have difficulty finding solutions and taking action towards solving everyday problems. The 6-step approach; Define the problem, Brainstorm solutions to the problem, Evaluate each solution, Select a solution to try, Try the solution, and Evaluate the solution (DBESTE) is reviewed through the use of worksheets. Appropriate planning strategies are then reviewed in conjunction with their new problem solving skills (e.g., You selected to call an exterminator for your roach problem, what steps do you have to take to accomplish this task?). Further, learning how to shift set or change their behaviors

when a previously identified "solution" is not working is also reviewed with real life examples (e.g., How many times do you tell yourself you will clean out the garage each weekend without success before you try another strategy? You still have a mouse infestation after trying your homemade remedy – when do you try something else?). Finally, basic organizational skills (e.g., filing, categorization) and maintenance routines (e.g., trash removal, home cleaning) are reviewed and operationalized. Daily homework is emphasized and reviewed at the beginning of each session.

Exposure to Discarding and Acquiring Modules (19 sessions in CREST condition; all 26 sessions in ET control condition). Symptoms of acquiring and saving are themselves avoidance behaviors that are performed to avoid internal distress related to negative thoughts and emotions[72]. Avoidance serves to reduce distress related to the beliefs regarding the necessity and utility of possessions[73,72]. In the CREST condition, the second part and the majority of treatment is dedicated to ET for discarding and not acquiring while in the control condition, the entire treatment will consist of ET. The ET utilizes in-vivo exposure exercises, beginning in the therapist's office, but eventually taking place during home visits as these home sessions assist with the generalization of their new skill sets[22]. Pure ET is different than exposures provided in CBT for HD in that ET is not combined with cognitive therapy, motivational interviewing or any of the other components of CBT for HD.

In ET for HD, the clients develop a hierarchy list of spaces in that evoke progressively greater fear if they were faced with having to make a decision about discarding. These discarding scenarios or spaces may range from some that are relatively easy to others that are incredibly difficult. For HD clients, fear hierarchies typically start with a space that has low clutter volume or there is less of an urge to save a particular type of item in that environment. The patient and therapist collaborate on creating this list and selecting a mild to moderately difficult space where the client will start exposure exercises.

Concrete rules are established for exposure activities, such as only two piles to place items in during exposure (keep and discard), only the client can make the decision about the item, only handle the item once, kept items must be put away immediately after the exposure, discarded items must go to the trash/recycle/donation area immediately after the exposure, and report distress ratings out loud to track progress.

Clients are asked to bring the items into the office from the identified space on their hierarchy and then to pick up each item one by one and make a choice about keeping and discarding. There is an agreed upon length of time that the patient must engage in exposure exercises and the rationale for ET is reviewed at the beginning of each session. During exposures, therapists remind patients of their self-identified treatment goals and concrete challenge questions (e.g., How does keeping this help you with your hoarding problem? Would you be able to live without this item? If you really need it again, could you get one?) are asked in an effort to encourage discarding. Clients progress through own hierarchy and move into different spaces or scenarios once the space is completed.

Ideally, habituation occurs over time and context and patients learn that they can tolerate the distress from the exposure activity[74]. Even if their distress does not decrease during the exposure, clients learn experientially that they can tolerate the distress. When obsession-related stimuli are repeatedly triggered, their toleration of fear increases[75]. Throughout sessions, clients learn that they do not experience terrible consequences when experiencing states of high anxiety while discarding. The goal is to generalize lessons learned from repeated exposures that trigger the feared reason for discarding or not acquiring.

A small percentage of clients have excessive acquiring behaviors. If it is determined excessive acquiring is a problem based on the initial assessment (SI-R, Acquiring Subscale; score >13), exposures to not acquiring are designed (e.g., go to Walmart with cash in hand and not purchase items, go to swap meet and not purchase items, watch home shopping network with credit card in hand and not purchase) and guided by a hierarchy. The clinician will work with the participant on these non-acquiring exposures and discarding exposures simultaneously.

Control Group: Exposure to discarding and acquiring (26 sessions). We propose to use a robust control condition, ET, with the same frequency and amount of therapist contact as CREST. Twenty-six weekly, individual ET sessions (6 months) will be delivered. The control group will receive ET for all 26 sessions and no cognitive training. As in CREST, the ET sessions will be manualized and copies utilized during session by both the patient and therapist.

While we considered using the currently available CBT for HD[21] as our control condition, ET will allow examination of the specific role that exposure plays in decreasing HD symptoms. Thus, we are investigating

the impact of a concentrated individual component of treatment, which we believe is the true mechanism of HD symptom change. Further, we may find subgroups that benefit from a specific condition, allowing for refinement of future treatment.

Staff

Two part-time PhD-level clinicians experienced in exposure-based treatments will be hired and trained in both CREST and ET. These clinicians will be at the postdoctoral level and license eligible. Study clinicians will rotate on taking new participants. Volunteer research assistants (VRAs) will be utilized as clinician extenders due to the necessity of at-home labor assistance required in cluttered homes. Our lab has a history of recruiting and maintaining VRAs for HD treatment studies and currently has 10 HD clinical VRAs that are approved by our local VA. For the proposed project, a total of 6 clinical VRAs will be trained (see below) along with the clinicians. The VRAs will attend home visits with the study therapists and assist with exposure activities. The VRAs are present during supervision and their work is audiotaped. Two part-time (15 hrs per week) independent raters will be hired to conduct in home assessments throughout the study. One part-time study coordinator/data manager will be hired for the project and assist with phone screening, data organization/entry, and recruitment.

Training and Supervision

Both the clinicians and VRAs will attend the initial training and weekly supervision. The amount of training and preparation proposed for this investigation is influenced by standard clinical practice within mental health care and consistent with many VA evidence based training roll-outs. The initial 2-day training workshop (8 hours per day) will be provided by Dr. Ayers and will include review of manuals and homework sheets, videotaped demonstration, and role-played/modeled delivery of CREST and ET. Dr. Ayers has conducted over 10 HD treatment workshops to mental health clinicians. Weekly 90-minute face-to-face skills based group supervision will be provided by Dr. Ayers. Supportive coaching, modeling, role-play, review of session audiotapes, and feedback of fidelity ratings will be used in supervision. The type of training program we have designed that places an emphasis on supervision and fidelity is also consistent with training models used for roll out of evidence based practices in large healthcare programs (e.g., VA). Independent raters will receive an initial 2-day training on neuropsychological assessment and ongoing weekly supervision from Dr. Elizabeth Twamley (Co-I).

Fidelity

In an effort to avoid therapist effects, the therapists will deliver both treatment conditions. While we considered issues of contamination, the cognitive rehabilitation components vary from the exposure sessions thus it would not be intuitive or feasible to deliver rehabilitation modules during the exposure sessions. The ET sessions start with the rationale for exposure and quickly moving to exposure to discarding and not acquiring; leaving no room for teaching compensatory cognitive strategies.

In the present protocol, we propose to promote and measure treatment fidelity in several ways. First, in addition to training outlined above, study therapists will receive weekly supervision from Dr. Ayers. Second, therapists will have a checklist to guide them during the conduct of sessions. Third, all sessions of both interventions will be videotaped, and a random sample of 20% of videos from each randomized group will be selected by study statistician, Dr. Golshan, and will be reviewed within a week and rated for treatment fidelity and discriminability by Dr. Ayers utilizing a standard checklist to document presence of all program elements (See Appendix 3 and 4). We will define therapist compliance as 90% adherence to the items on the weekly checklist. Specific plans have been developed for responding to sub-threshold adherence, ranging from extra training to statistically controlling for adherence to replacing therapists (in the extreme).

With respect to our blind, independent raters, intra-class correlation coefficients (ICC) will used to determine inter-rater reliability of individual items and scale scores for appropriate measurements. We will establish a good inter-rater reliability (ICC=.85) prior to any data collection. Each rater's score will be compared to the "Gold Standard" rater. The raters are required to achieve an ICC of at least .85 for all rating scales. To protect against "rater drift" over the course of the study, all raters will undergo reassessment semiannually. New raters hired during the course of the study will go through the reliability training and testing described above before they rate subjects.

Measures

Demographic and Diagnostic assessment (at baseline only) will assess inclusion criteria, including diagnosis, co-morbidities and cognitive impairment (Table 2). Diagnosis of HD will be assessed using the Structured Interview for Hoarding Disorder (SIHD)[67], a recently developed but well-validated clinician-administered instrument for diagnosing HD based on the DSM-5 criteria. The SIHD has demonstrated strong inter-rater reliability, convergent validity, and discriminant validity. Potential comorbidities will be assessed with the Mini-International Neuropsychiatric Interview (M.I.N.I.)[68], a brief diagnostic interview that uses decision-tree logic to assess different mental disorders. This interview has demonstrated excellent reliability and validity (it produces the same diagnoses as the Structured Clinical Interview for DSM-IV Axis I Disorders (SCID-I) or Composite International Diagnostic Interview (CIDI) in 85-95% of cases), and it can be administered much more rapidly than other diagnostic interviews (e.g., SCID-I).

Hoarding symptom severity (primary outcome) will be measured using the Savings Inventory-Revised (SI-R)[56], a 23-item self-report measure used to assess common hoarding symptoms. Subtests include excessive clutter, compulsive acquisition, and difficulty discarding. The SI-R has demonstrated good internal consistency, divergent validity, concurrent validity, divergent validity, test-retest reliability in clinical samples with hoarding. The total score will be used for analyses.

Functional Capacity (secondary outcome) will be assessed using two different measures. The UCSD Performance-Based Skills Assessment (UPSA)[76] is an assessment of everyday functioning skills involved in household chores (e.g., writing a shopping list based on a provided recipe), communication (e.g., rescheduling a doctor's appointment), finance (e.g., paying a utility bill), recreation planning (e.g., planning an outing to the beach or zoo), and transportation (e.g., reading comprehension of a bus schedule). The UPSA has demonstrated high inter-rater reliability (0.91) and convergent validity with other performance-based measures. The total UPSA summary score will be used for analyses. The UCSD SORT Test (U-SORT)[77] will be used to measure Veteran participants' organizational skills as they relate to functional capacity. During the administration of the U-SORT, participants are instructed to sort 42 household objects (e.g., bent and unbent paper clips, used and unused condiment packets) from a hypothetical "junk drawer" into either "keep" or "trash" piles. Participants are given two minutes to complete the task and one point is awarded for each correctly sorted item, for a total of 42 points. The U-SORT has high internal consistency (α = .86) and adequate convergent validity. The total U-SORT score will be used in analyses.

Self-reported functioning (secondary outcome) will be assessed with the Specific Levels of Functioning test (SLOF)[78], a 43-item questionnaire regarding areas such as interpersonal relationships, participation in community activities, and work skills. The SLOF has demonstrated excellent reliability and internal consistency. Further, the World Health Organization Disability Assessment Schedule (WHODAS 2.0)[79], a 36-item, six domain (Cognitive, Mobility, Self-Care, Getting Along, Household, Work, Participation) assessment instrument developed by World Health Organization (WHO) to provide a standardized method for measuring health and disability across cultures will also be utilized. The WHODAS demonstrated total internal consistency (0.96) as well as domain-specific reliability (0.79-0.98) and concurrent validity with similar disability measures. A summary score will be generated using the "complex" scoring method.

Quality of Life (secondary outcome) will be assessed using the Quality of Life in Neurological Disorders (Neuro-QoL) Positive Affect and Well-Being Short form[80]. The Neuro-QoL Positive Affect and Well-Being - Short Form is a 9-item self-report measure that assesses aspects that related to a sense of well-being, life satisfaction, purpose and learning. The total score will be used for analyses.

Avoidance (intermediary target) will be measured by the Acceptance and Action Questionnaire 2 (AAQ-2)[81], a 10 item, self-report measure of experiential avoidance and immobility that has demonstrated strong reliability (0.78-0.88). There is a precedence of using the AAQ-2 for measuring avoidance as a mediator of HD[82]. A Behavioral Avoidance Test (BAT) will also be utilized for assessing avoidance to sorting and discarding possessions in HD patients. The BAT has been validated for measuring avoidance in clinical obsessive compulsive disorder (OCD) samples[83], and will be adjusted for use in HD samples. Participants and assessors will select three kinds of personal objects that cause the Veteran participants distress to discard and agree on 3-7 steps that would go into sorting/discarding the object (deciding whether to keep it, throw it away, or give it away; deciding where to put the object if keeping it, where to throw it away if tossing it, and where to take it if giving it away, etc.). The assessor will observe the participant performing as many steps for each

object as they feel comfortable doing and for each step the participant at least partially completes, they will give a rating of their Subjective Units of Distress (SUDS) from 0-10. For each of the three object tasks the clinician will give a rating of 0-2 of the avoidance they observe in the participant, where zero indicates no avoidance, one indicates partial avoidance or that the participant did not fulfill all of the steps of the task, and two indicates that there was complete avoidance of the entire task of discarding the object. This process will be repeated in three rooms of their home; ideally the kitchen, bedroom, and living room pending accessibility. To create a composite, score the SUDS and the Avoidance scores will each be divided by their respective standard deviations and then summed together.

Executive Functioning (intermediary target) will be assessed using multiple measures. The Delis-Kaplan Executive Function System (D-KEFS)[65] was specifically designed to detect deficits in high-level cognitive functions, such as from mild brain damage to the frontal lobe[84] and has demonstrated strong reliability and validity in various clinical populations. We will administer five tests from the D-KEFS that examine components of executive functioning, such as inhibition (the Color-Word Interference Test), cognitive flexibility or switching (Verbal Fluency, Trails Tests, Design Fluency) and Planning (Towers). The Penn Conditional Exclusion Test (PCET)[85] will be used as a measure of executive functioning in which Veteran participants must perform a computerized sorting task with three shifting sorting principles. The PCET is similar to the Wisconsin Card Sorting Test[86], a classic test of executive functioning, but it has four alternate test versions that can be administered with comparable validity and reliability, better allowing for repeated testing over time.

Table 2. Assessments and Timeline

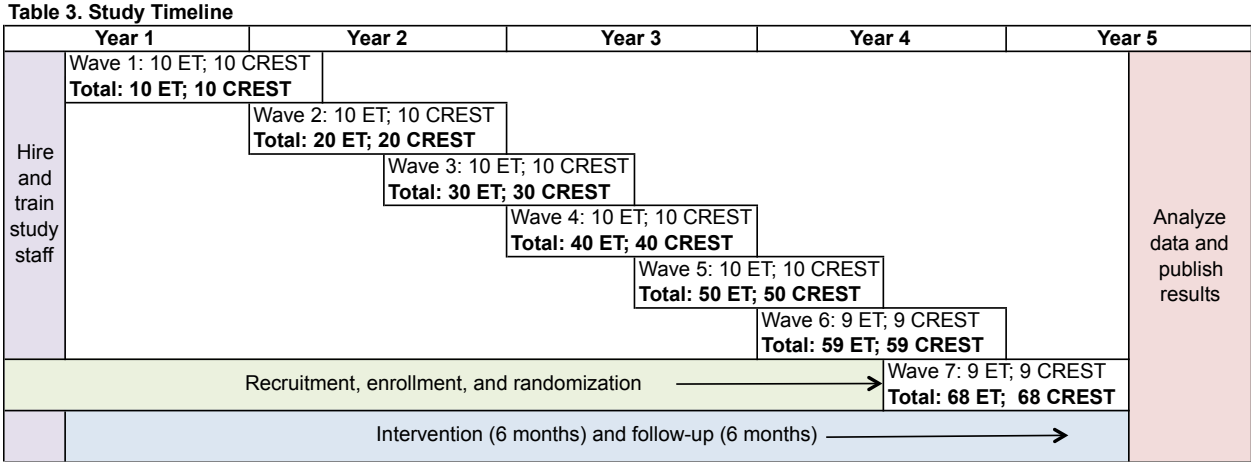
Measure	Baseline	Session 7	Session 13	Session 21	Post Assessment (After Session 26)	3-Month Follow-up	6-Month Follow-up
Diagnostic Assessment							
Structured Interview for Hoarding Disorder (SIHD)	X						
Mini-International Neuropsychiatric Exam (MINI)	X						
Hoarding severity (Primary Outcome)							
Savings Inventory-Revised (SI-R)	X	X	X	X	X	X	X
Functional Capacity (Secondary Outcome)							
UCSD Performance-Based Skills Assessment (UPSA)	X	X	X	X	X	X	X
UCSD SORT Test (U-SORT)	X	X	X	X	X	X	X
Self-Reported Functioning (Secondary Outcome)							
Specific Levels of Functioning test (SLOF)	X	X	X	X	X	X	X
World Health Organization Disability Assessment Schedule (WHODAS 2.0)	X	X	X	X	X	X	X
Quality of Life (secondary outcome)							
Quality of Life in Neurological Disorders (Neuro-QoL)	X	X	X	X	X	X	X
Executive Functioning (Intermediary Target)							
Delis-Kaplan Executive Function System (D-KEFS) Color-Word Interference, Verbal Fluency, Trails Tests, Design Fluency, Towers	X	X			X	X	X
Penn Conditional Exclusion Test (PCET)	X	X			X	X	X
Avoidance (Intermediary Target)							
Behavioral Avoidance Test (BAT)	X	X	X	X	X	X	X
Acceptance and Action Questionnaire 2 (AAQ-2)	X	X	X	X	X	X	X

Adherence to treatment (mediator) will be measured in the following ways: (1) Number of sessions attended; (2) Therapist ratings of homework compliance (0-100%) based on written assignment review and patient report after each session. Composite adherence score will be calculated on a 10-point Likert scale ranging from 0 (Almost never/never adherent) to 10 (Always adherent).

Timeline

The first 3 months of the project are designed to hire staff, train clinicians, the blind raters, and VRAs (Table 3). There will be a total of 7 overlapping study waves, each lasting 12 months, in which we will enroll 9-10 patients. During the last 3 months of the project, we will finalize data analysis, publish the results and determine the next step for competitive funding.

Table 3. Study Timeline



Statistical Plan

Missing data will be minimized by intensive training of the raters in techniques of clarifying answers and checking questionnaires while Veteran participants are on-site. When missing values are identified, several approaches to acquire the necessary data will be employed: (1) if at all possible, Veteran participants will be rescheduled within 24 hours of completion of tests or interviews; (2) if a participant cannot be rescheduled, an effort will be made to send a tester to the individual's place of residence within the same period; (3) missing data will be examined to assess randomness. Missing data (i.e. loss to follow-up) will be tested to determine if it is informative, and the methods developed by Diggle[87] and Ridout[88] to test for completely random dropouts will be applied. If necessary, the techniques of Paik[89] applicable to generalized estimating equations models will be used to impute missing values. However, the primary data analysis method allows inclusion of subjects with missing data or those who terminated the study early, without relying on data imputation procedures. Furthermore, we will test whether the drop-outs are random or systematic by comparing the drop-outs with the study completers on the baseline data. The importance of missing data will be examined in the second stage of data analysis (see below).

Data analysis will be performed in multiple stages. In the initial stage, preliminary analyses will be performed by examination of the distribution of variables to assess their characteristics (means, standard deviations, and skewness). Descriptive statistics and graphs will used to summarize the characteristics of the study population. Continuous measures will be tested for normality and homogeneity of variance. Non-normally distributed variables will be transformed to meet the normal distribution assumption for linear models. Randomization will be assessed by performing a series of Wilcoxon-rank sum tests, Chi-square or Fisher's exact tests to compare the groups on demographic and initial baseline clinical variables. Any variables on which the groups differ initially will be explored as covariates in subsequent analyses, and we will take this into account in the interpretation of the outcome. At the second stage, we will perform sensitivity analyses for the impact that potentially informative missing data may have on the analyses. We will use pattern-mixture models to assess if there is bias due to drop out or missing data. In addition, we will consider the extension of the pattern-mixture models as described by Guo, Ratcliffe, and Ten[90], which includes the incorporation of random effects in the pattern-mixture model, which allow subject-to-subject heterogeneity.

At the final stage, hypotheses will be tested using Mixed Effects models, a generalized linear model described by Gibbons et al.[91], Hedeker et al.[92] and Laird et al.[93]. Responses at baseline, throughout treatment (sessions 7, 13, 21), post treatment (session 26), 3- and 6-month follow-up will be nested within subjects. Intercept and slope will be modeled as random effects nested within subject; Treatment group, time, and group-by-time interaction will be a fixed effect. A fully saturated treatment by time model will be utilized for inference. Co-variance structure will be chosen based on Akaike's Information Criterion (AIC). Data will be analyzed from all randomized subjects on whom we have a baseline assessment and at least one post-baseline evaluation. If it is needed based on analyses on stage one, baseline covariates (e.g., demographics) will also be added to the model. All hypothesis tests will be intent-to-treat, and will be two-sided at the .05 level. We will adjust p-values for multiple testing within each hypothesis separately by using a Bonferroni correction whenever more than one variable is being tested. Although we didn't have any drop-out in out pilot study, we included a drop-out rate of 10% after the baseline assessment but prior to the first post-baseline evaluation in our power analyses.

Hypothesis 1.1-1.3: Relative to those assigned to ET alone, CREST participants will show significant reduction in **hoarding severity** (primary outcome), improvement in **functioning** (both functional capacity and self-reported functioning; secondary outcomes) and improvement in **quality of life** over time.

Independent Variables: Treatment Groups (CREST and ET), Time points (baseline, throughout treatment (sessions 7, 13, 21), post-treatment (session 26), 3 and 6-month follow-up).
Dependent Variable: Hoarding severity (SI-R), Functioning (UPSA, USORT, SLOF, WHODAS 2.0.), Quality of Life (Neuro-QOL).
Statistical Analysis: The dependent variable will be analyzed by mixed effects model methods as described above. Baseline demographic and other clinically important characteristics at baseline will be assessed for imbalance between the two groups and their association with the outcome using a univariate analysis. Those found to be significantly unbalanced between intervention groups ($p < 0.10$) and moderately associated with the outcome ($p < 0.15$) will be included as covariates in the model. We will adjust p-values for multiple testing within each hypothesis separately by using a Bonferroni correction whenever more than one variable is being tested

Hypotheses 2.1: Treatment adherence, improvement in executive functioning, and reduction in avoidance will mediate hoarding symptom reduction in CREST.

Dependent Variables: Hoarding severity (SI-R)
Independent Variables: Executive Functioning (D-KEFS scaled scores and PCET), Avoidance (AAQ and BAT) and Treatment Adherence score.
Statistical Analysis: The effect of mediator variables will be tested following the procedures recommended by Kraemer[94], Bauman[95], Cohen and Cohen[96], and Busemeyer and Jones[97]. Briefly, treatment mediators explain how or why one variable affects the treatment effect and are defined as mechanisms through which a treatment might achieve its effects. The mediator's effect is measured during the treatment and is correlated with treatment choice and may possibly be a result of treatment and have either a main or interactive effect on the outcome. The dependent or outcome measure (SI-R) will be evaluated throughout the study and will be used as a repeated measure for all the available time points. In essence, the slope of each outcome measure on time will be used as a dependent measure. The mediation effect will be quantified by calculating the proportion of the treatment effect explained by the mediator and bootstrap technique will be used to calculate a 95% confidence interval for mediation effect. Random effects models will be used for analysis and the change score of mediators will be included in the model.

Hypotheses 2.2: Symptom severity reduction and improved functional capacity will mediate improved quality of life in CREST.

Dependent Variables: Hoarding severity (SI-R), Functioning (UPSA, USORT, SLOF, WHODAS 2.0)
Independent Variables: Quality of life (Neuro-QoL)
Statistical Analysis: The effect of mediator variables will be tested similar to the hypothesis 2.1.

Exploratory Aim 3: To explore **age and baseline executive functioning** as moderators in order to determine subpopulations in which CREST work best.
Exploratory Aim 4: To examine time to **maximum effect of treatment gains.**

These aims are focused on exploring on whom and under what conditions the treatment works best. These analyses may identify subpopulations with possibly different mechanisms and must be a baseline or pre-randomization characteristic that could impact treatment outcome. We will focus on the moderator effect of age and baseline executive functioning and their interactions with treatment group. Age will be explored as continuous and ordinal (i.e. 45-55, 56-65, 66-75, 76-85) in format. If the interaction effect is significant, pairwise comparisons of groups will be performed. These analyses will examine interactions of group with time at each level of the moderator, and evaluate treatment within each moderator subgroup. These exploratory aims will be examined using random effects regression analyses for each candidate moderator as a main effect and in interaction, similarly to methodology used for the hypothesis testing. Furthermore, we will explore time to maximum effect of treatment gains (>35% reduction on SI-R) using survival analysis. The estimated survivorship curves will be obtained from Kaplan-Meier maximum likelihood estimates for each treatment group. The parameter estimates and 95% CI for the hazard ratio will be obtained from the Cox proportional hazards. Subjects who do not respond to treatment will be censored on their last visit date.

Power Analysis
Power analyses are performed for all hypotheses, incorporating the longitudinal nature of the design. Within this design, repeated observations within subjects are potentially correlated. This has a profound impact on the resulting tests of significance[98]. When the within subject correlation is properly incorporated, the repeated measures analysis takes full advantage of all information obtained from each subject, thereby greatly increasing statistical power over methods that compare treatments cross sectionally[99]. Using the method described by Diggle, Liang, & Zeger[100] and Ahn et al.[101], power calculations can be made for repeated measures designs under specified assumptions. Hedeker, Gibbons, & Waternaux[102] extend Diggle et al.[1-3]'s method for various covariance structures. Procedures describes by Hedeker et al.[102] for Random Regression Models which incorporated these principles was used for the proposed study to estimated needed sample size (RMASS program provided by Hedeker for the Random Regression Model). The design assumptions consist of a 2 treatment groups and up to 7 visits (baseline, sessions 7, 13, 21, 26, and 3- and 6-month follow-up) where the slope is the dependent measure. Other assumptions used were alpha-level, nature of the hypothesis (two-sided versus one-sided), drop-out rate, and variance-covariance matrix of the longitudinal data. We assume a Type-I error level (alpha-level) of .01 to 0.05, drop-out rate of 15% and an autoregressive covariance structure. Data reported in the preliminary studies section which included our own pilot studies were used to calculate effect sizes. From the preliminary analysis of pilot study, the estimated standard deviation (SD) for SI-R score ranges from 12.75-12.82 and the between visit correlation coefficient of 0.71. We considered a range of estimates of stand deviations of 12.5, 12.8 and 13.0, and correlations of .45 to 0.70 in our power analysis. We have reported large effect sizes of 1.02 and 1.51 from our previous studies which converted to differences of 60 to 37 for our primary outcome. However, for these calculations, we conservatively selected smallest clinically meaningful differences that we would like to be able to detect. This was then converted to a medium effect size, defined mathematically as a between-group difference increasing linearly from 0 at baseline to .5 SD units at the last time point. Calculations using the above assumptions indicated that the study will have minimum power of 80% to yield a statistically significant result for a medium effect size with the proposed sample size of 136 patients (68 for each of the two groups).