



Published in final edited form as:

*J Psychopathol Behav Assess.* 2019 March ; 41(1): 135–143. doi:10.1007/s10862-018-9701-7.

## Validation of a Behavioral Measure of Acquiring and Discarding in Hoarding Disorder

Hannah C. Levy<sup>1</sup>, Michael C. Stevens<sup>2</sup>, and David F. Tolin<sup>1,3</sup>

<sup>1</sup>Anxiety Disorders Center, Institute of Living, Hartford Hospital 200 Retreat Avenue, Hartford, CT 06106

<sup>2</sup>Olin Neuropsychiatry Research Center, Institute of Living, Hartford Hospital 200 Retreat Avenue, Hartford, CT 06106

<sup>3</sup>Department of Psychiatry, Yale University School of Medicine

### Abstract

Behavioral assessment is an important component of evidence-based assessment and treatment in anxiety and related disorders. The purpose of the current study was to validate a behavioral measure of difficulty discarding and acquiring, the core features of hoarding disorder (HD). Seventy-eight patients with a primary diagnosis of HD completed a computerized acquiring and discarding task; the task consisted of making simulated decisions about acquiring and discarding items of varying monetary value. A subset of patients ( $n = 42$ ) went on to receive cognitive behavioral therapy (CBT) for HD and completed the computer tasks again after treatment. An additional 30 age- and sex-matched healthy control participants (HCs) also completed the acquiring and discarding tasks. Results showed that HD patients saved and acquired more items than the HC group, and had longer response times during the tasks. In support of the convergent validity of the tasks, item decisions and reaction times were positively correlated with established measures of HD symptoms. Among treatment completers, items saved and acquired and response times decreased from pre- to post-CBT, suggesting that the tasks were sensitive to detect treatment-related changes in difficulty discarding and acquiring behaviors. The findings support the validity of the discarding and acquiring tasks in measuring HD symptoms, and are discussed in terms of the potential advantages of behavioral measures in HD treatment and research.

---

Correspondence concerning this article should be addressed to Hannah C. Levy, Anxiety Disorders Center, Institute of Living/Hartford Hospital, 200 Retreat Avenue, Hartford, CT 06106. Tel.: 860-545-7847. Fax: 860-545-7105., hannah.levy@hhchealth.org.

**Ethical approval:** All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Compliance with Ethical Standards

1. Disclosure of potential conflicts of interest:

b. Conflicts of interest: Authors Hannah Levy, David Tolin, and Michael Stevens declare that they have no conflicts of interest.

2. Research involving human participants:

a. As reported in the text, the study was approved by the local Institutional Review Board. The statement about ethical approval has been included above, as instructed.

3. Informed consent:

a. As reported in the text, participants provided informed consent prior to participating in the study.

## Keywords

hoarding disorder; acquiring; discarding; behavioral assessment

---

Hoarding disorder (HD) is characterized by difficulty discarding personal possessions, regardless of their value (American Psychiatric Association, 2013). Other features of HD include compulsive acquiring and difficulties with decision-making (Steketee & Frost, 2003; Tolin, 2011). Among 653 self-identified hoarding patients who completed an online survey, 558 (86%) indicated at least moderate acquiring problems (Frost, Tolin, Steketee, Fitch, & SelboBruns, 2009), suggesting that compulsive acquiring is a common symptom of HD. Nevertheless, DSM-5 included excessive acquiring as a diagnostic specifier rather than a formal criterion required for HD diagnosis, possibly due to the fact that not all HD patients endorse excessive acquiring. Current neurobiological models of HD implicate difficulties with decision-making as underlying mechanisms of the disorder (Tolin, Stevens, Villavicencio, et al., 2012). Specific decision-making impairments that have been observed in HD include categorization (Grisham, Norberg, Williams, Certoma, & Kadib, 2010), self-reported fears and avoidance of decision-making (Frost, Tolin, Steketee, & Oh, 2011), and set shifting (Ayers et al., 2013), even when changing set would result in rewards and/or minimize punishments (Lawrence et al., 2006).

Currently, measures of difficulty discarding and acquiring in HD are limited to self-report questionnaires (e.g., the Saving Inventory-Revised, SI-R; Frost, Steketee, & Grisham, 2004) and one clinician-administered interview (the Hoarding Rating Scale-Interview, HRS-I; Tolin, Frost, & Steketee, 2010), which may not accurately capture symptom severity. Indeed, HD patients may lack insight into the severity of their symptoms (Worden, DiLoreto, & Tolin, 2014), raising questions about their ability to accurately reflect their current symptoms, either by self-report or during a diagnostic interview (DiMauro, Tolin, Frost, & Steketee, 2013). Similarly, although self-report measures of decision-making difficulties exist (e.g., Frost Indecivness Scale, FIS; Frost & Shows, 1993), HD patients may underestimate their cognitive abilities (Moshier et al., 2016), again calling into question the accuracy of self-report measures for this population. With these limitations of current measures in mind, the aim of the present study was to validate a behavioral measure of HD symptoms, which assesses acquiring and discarding decisions in real time. The measure requires participants to make simulated decisions about acquiring and discarding possessions while recording reaction time (response latency), which is used as a proxy of decision-making difficulty.

It is clear that more direct measures of target behaviors and symptoms in HD are needed. Behavioral measures are important (yet underutilized) in treatment outcome research (e.g., Haynes, Keawe'aimoku Kaholokula, & Yoshioka, 2008); in particular, these measures may be more specific and narrowly focused than are more general measures of psychopathology (e.g., measures of personality traits, Global Assessment of Functioning scores), thus minimizing error variance and alternative explanations of observed change during treatment. Direct measures of disorder-relevant behaviors have improved assessment approaches in HD-related disorders, such as obsessive-compulsive disorder (OCD) and anxiety disorders.

For instance, behavioral approach tests, which assess approach behavior towards feared stimuli, are commonly used to assess anxiety and phobic symptoms. One example of a behavioral approach test was developed by Cogle, Wolitzky-Taylor, Lee, and Telch (2007) and validated by Najmi, Tobin, and Amir (2012) to assess contamination-related fear and avoidance, which are commonly observed in patients with OCD (Rachman, 2004; Ruscio, Stein, Chiu, & Kessler, 2010). Another example is the use of live speech tasks and/or video feedback when assessing social and performance anxiety; these measures provide objective data about anxiety symptoms (e.g., stuttering, blushing). Socially anxious patients may overestimate how noticeable these symptoms are, making it important for clinicians to observe them live. Importantly, behavioral measures provide real time information about stimulus approach and avoidance, which has important implications for assessment and treatment planning. Similarly, a direct behavioral measure of HD symptoms would provide useful information about approach and avoidance behavior in the context of acquiring and discarding which, when used in combination with other assessment tools such as diagnostic interviews and self-report questionnaires, would provide a comprehensive picture of symptom severity.

To this end, in our earlier research we utilized a paper discarding task, in which HD patients and healthy controls (HCs) sorted and discarded their own and the experimenter's junk mail in real time (Tolin, Kiehl, Worhunsky, Book, & Maltby, 2009; Tolin, Stevens, Villavicencio, et al., 2012). Although the percentage of personal vs. experimenter-owned junk mail items discarded was positively correlated with difficulty discarding on a standardized interview of HD symptom severity, indicating convergent validity of the task with related measures, this discarding task had several important limitations. First, it is impractical to require that research participants or clinical patients bring in their own mail to the laboratory or clinic. Second, the task only assessed junk mail discarding, which may not generalize to discarding of other personal possessions. Third, the task was limited to discarding and did not assess acquiring behavior.

A computer-based acquiring and discarding task, developed by Preston et al. (2009), addressed some of these limitations. In this task, nonclinical undergraduate participants were presented with pictures of a large number of objects, varying in usefulness and monetary value, and asked to choose which items they would (hypothetically) want to take home for free. Participants then completed several rounds of simulated discarding of the items they had "acquired," each round increasing in pressure to discard more items. In support of the convergent validity of the task, high acquirers showed elevated scores on HD-related self-report measures relative to participants who acquired less items during the task. Other strengths of this task are feasibility and easy administration, assessment of both acquiring and discarding behavior, and inclusion of a variety of items ranging in value. On the other hand, the use of a nonclinical undergraduate sample limits the generalizability of the findings to clinical patients with HD. Further, the task was not evaluated in a treatment context, so it is unclear whether the task is sensitive to treatment-related changes in hoarding symptoms.

To address these limitations, we administered a revised version of the Preston et al. (2009) task to a small pilot sample of HD patients ( $n = 6$ ) and healthy control participants ( $n = 6$ ;

Tolin, Stevens, Nave, Villavicencio, & Morrison, 2012). HD patients completed the task before and after receiving cognitive-behavioral therapy (CBT). Results showed that HD patients saved more items than did HCs during the discarding task, but the groups did not differ in terms of number of items acquired, reaction time when making acquiring decisions, or reaction time when making discarding decisions. Some task measures improved from pre- to post-treatment in the HD group, suggesting that the task was sensitive to treatment-related changes in hoarding behaviors. Of course, the small sample size makes these results tentative; replication in a larger validation study is needed.

To this end, the aim of the present study was to examine the psychometric properties of the discarding and acquiring tasks in a larger sample of HD patients and HCs. We assessed construct-related validity in relation to HD diagnosis and convergent validity with both clinician-administered and self-report measures of HD severity. We predicted that HD patients would save and acquire more items than would HCs. Based on our pilot work using the same task (Tolin, Stevens, Nave, et al., 2012), we expected that the discarding task would better discriminate between HD patients and HCs than the acquiring task (i.e., effect sizes for group comparisons would be larger for discarding than acquiring). Consistent with neurobiological models of HD (Tolin, Stevens, Villavicencio, et al., 2012), we predicted that HD patients would show significantly longer reaction times on the task than would HCs. We further predicted an interaction between group (HD vs. HC) and task (experimental, discarding or acquiring; vs. control, making decisions about non HD-specific control items) on reaction time, such that HD patients would demonstrate slower reaction times than HCs specifically for the experimental tasks. We also predicted that the acquiring and discarding tasks would correlate positively with clinician- and self-reported measures of HD. We also assessed changes in task measures during treatment, and predicted that the number of acquired and saved items and decision reaction time would decrease from pre- to post-CBT.

## Method

### Participants

Participants were 78 adults (age 20–65) with a primary diagnosis of HD of at least moderate severity who completed an intake assessment as part of a waitlist-controlled trial of CBT for HD (“HD group”). Those on psychiatric medications had to be on a stable dose for at least eight weeks and maintain the same dose for the duration of the study. Only antidepressants, stimulants, and benzodiazepines were permitted. Because the discarding and acquiring tasks were administered during functional magnetic resonance imaging (fMRI), a 24-hour washout of stimulant and benzodiazepine medications was required to reduce acute neural changes associated with these medications. Exclusion criteria for the HD group were current suicidality, psychosis, bipolar disorder, substance use disorder, any severe psychiatric problem requiring a higher level of care, and prior CBT for HD. A total of 130 potential participants were evaluated for study inclusion; 52 were excluded for not meeting the study eligibility criteria. The most common reasons for exclusion were not meeting criteria for HD ( $n = 14$ ), primary diagnosis other than HD ( $n = 10$ ), and substance use disorder ( $n = 8$ ).

An additional 30 age- and sex-matched healthy controls (“HC group”) were assessed. Participants in the HC group could not have any current or past psychiatric disorder.

Exclusion criteria for both groups were lack of English fluency; cognitive dysfunction that could interfere with the capacity to understand study procedures and/or provide informed consent; and history of anoxic or traumatic brain injury with loss of consciousness for more than five minutes. A total of 46 potential HC participants were evaluated for study inclusion; 16 were excluded for not meeting the study eligibility criteria. The most common reasons for exclusion in the HC group were current psychiatric disorder ( $n = 6$ ), abnormal MRI findings ( $n = 4$ ), and current HD symptoms ( $n = 3$ ).

## Measures

**Clinician-administered interviews.**—The *Diagnostic Interview for Anxiety, Mood, and Obsessive-Compulsive and Related Neuropsychiatric Disorders* (DIAMOND; Tolin et al., 2018) was used to determine participants' diagnoses. The DIAMOND is a structured diagnostic interview based on the DSM-5 that has demonstrated good reliability and validity estimates for anxiety, obsessive-compulsive, and depressive disorders, including HD (Tolin et al., 2018). We did not have inter-rater reliability data available for the DIAMOND. A modified version of the *Clinical Global Impression* (CGI) scale (Guy, 1976), called the CGI-HD (Tolin et al., in press), was used to determine overall HD symptom severity on six dimensions, including clutter, difficulty discarding, acquiring, health or safety hazard, functional impairment, and distress. Whichever domain yields the highest severity rating becomes the CGI-HD score (e.g., if clutter receives a rating of "severe" and all remaining domains receive lower severity ratings, the CGIHD score will be "severe"). In a prior study using the same data set, we found adequate interrater reliability for the CGI-HD (ICC = 0.72; Tolin et al., in press). The *Hoarding Rating Scale – Interview* (HRS-I; Tolin, Frost, & Steketee, 2010) was used to evaluate severity of hoarding symptoms. The HRS-I is a 5-item structured interview that assesses the core symptoms of HD (difficulty discarding, acquiring, and clutter) as well as distress and impairment associated with these symptoms. The HRS-I has been shown to effectively discriminate between HD and nonHD samples (Tolin et al., 2010), and showed excellent internal consistency in the current sample ( $\alpha = 0.97$ ).

Interviewers were psychology postdoctoral fellows or licensed psychologists who received extensive training in administration of the DIAMOND and HRS-I. Interviewers met regularly with the principal investigator of the study (a licensed psychologist with expertise in assessment and treatment of HD) to discuss assessment results and resolve any questions. Interviews were audio-recorded and coded by independent raters to establish inter-rater reliability.

**Self-report measures.**—The *Saving Inventory-Revised* (SI-R; Frost et al., 2004) was used to assess hoarding severity. The SI-R contains three subscales (acquiring, difficulty discarding, and clutter) corresponding to the three core features of HD. Items are rated on a 5-point Likert scale (0 = *none* and 4 = *almost all/complete*) with higher scores indicating greater hoarding severity. The SI-R has demonstrated adequate reliability as well as strong convergent and discriminant validity (Frost et al., 2004), and it showed excellent internal consistency in the current sample (SI-R total score,  $\alpha = 0.98$ ; acquiring subscale,  $\alpha = 0.94$ ; difficulty discarding subscale,  $\alpha = 0.95$ ; clutter subscale,  $\alpha = 0.98$ ). The *Frost Indecisiveness Scale* (FIS; Frost & Shows, 1993) was used to assess difficulties with decision-making. The

FIS has two subscales, the first measuring fears about decision-making (e.g., “I often worry about making the wrong choice”) and the second assessing positive attitudes about decision-making (e.g., “Once I make a decision, I stop worrying about it”). The FIS has demonstrated adequate reliability in undergraduate (Frost & Shows, 1993) and clinical (Frost et al., 2011) samples; it showed excellent internal consistency in the current sample (fear subscale,  $\alpha = 0.92$ ; positive attitudes subscale,  $\alpha = 0.87$ ). The SI-R and FIS were administered prior to the discarding and acquiring tasks.

**Computer tasks.**—The discarding and acquiring tasks were adapted from Preston et al. (2009) and tested in our pilot research (Tolin, Stevens, Nave, et al., 2012). We made three modifications to the original tasks: 1) to increase pressure, we limited the number of “acquired” and “saved” items to only those that could fit in a standard shopping cart; 2) to reduce fatigue effects, we presented different stimuli during the discarding part of the task, rather than asking participants to discard the items they had just “acquired” during the acquiring phase of the task; and 3) for comparison purposes, we added a control task (see below). In the current study, the discarding and acquiring tasks were administered while participants completed fMRI as part of a larger study examining changes in brain activation during CBT whose results will be described in a separate report. The computer tasks were displayed on a computer screen behind the participants’ heads. Participants viewed the screen using a mirror attached to the head coil of the fMRI scanner. Participants were instructed to lie still and indicate their responses via a button press. They were given 7.5 seconds to make a response; if no response was made, the next stimulus was presented. Both tasks consisted of 30 test items and 30 control items, which were presented at random. Prior to beginning the tasks, participants completed several practice items. Immediately following the discarding and acquiring tasks, participants provided subjective fear/anxiety and sadness/regret ratings via a visual analogue scale from 0 (*Not at all*) to 100 (*Extremely*).

**Discarding task.**: Participants were presented with pictures of household items that were being sorted during the discarding task. They had to decide whether to keep or discard each item. Item decisions, as well as reaction time, were recorded and used for analyses. The task was introduced as follows: “In this task, you will see a series of pictures. Imagine that all of the items pictured belong to you. Imagine that you are doing some spring cleaning. You can decide which items you would like to keep, and which items you would like to discard. Discarding means the item will go in a trash and be taken away by the trash collectors. You will not be able to recycle, sell, or donate any of the items. Underneath each picture you will see the words ‘KEEP IT’ on the left, and ‘DISCARD IT’ on the right. If you want to KEEP an item press the button under your index finger. If you want to DISCARD an item press the button under your middle finger. When you KEEP an item, a green border will appear around the picture, meaning it is still YOURS. When you DISCARD an item, a red border will appear around the picture, meaning you have put it into the trash. There is only one catch. You can keep as many things as you want; however, you can only keep as many things as will fit in a standard shopping cart like this one [a picture of a shopping cart was shown]. Once you have filled your shopping cart, you will not be able to add or remove any items. Therefore, select your items carefully, only keeping items you really want. Remember, the



items you see belong to you. They are in your home, and you can keep or discard whichever items you want.”

**Acquiring task.:** In this task, participants were presented with pictures of items for sale. They had to decide whether to acquire or leave each item, and told that they could only acquire enough items to fill a standard shopping cart. As above, item decisions and reaction time were recorded for analyses. This task was introduced as follows: “In this task, you will see a series of pictures. Imagine that all of the items pictured are here today in this building. You can take home anything you want today, for free. Underneath each picture you will see the words ‘TAKE IT’ on the left, and ‘LEAVE IT’ on the right. If you want to TAKE an item home today, press the button under your index finger. If you want to LEAVE an item here today press the button under your middle finger. When you TAKE an item, a green border will appear around the picture, meaning it is now YOURS to take home. When you LEAVE an item, a red border will appear around the picture, meaning you will not be taking it home. There is only one catch. You can take home anything you want today, for free; however, you can only take as many things as will fit in a standard shopping cart like this one [a picture of a shopping cart was shown]. Once your shopping cart is full you will not be able to add or remove any items. Therefore, select your items carefully, only taking things that you really want. Remember, the items you see do not belong to you yet. But you can take home anything you want today, for free.”

**Control task.:** For comparison, we included a control task that was not specific to HD symptoms or behaviors. In this task, participants were presented with pictures of various objects and asked to decide whether the object was alive (either currently or in the past, such as an apple) or never alive (e.g., a plastic toy). Again, item decisions and reaction time were recorded. The control task was introduced immediately following the instructions for the two active tasks, as follows: “There also is another task you have to do. For some groups of items, you don’t get to choose whether or not take them home. Instead, we will ask you to look at each item and indicate if it is ‘Once Alive’ or ‘Never Alive.’ ‘Once Alive’ means that something in the item is alive, or it is made of materials that were once alive. For example, items made of paper, cloth, or things like food. These are sometimes thought of as ‘organic’ or coming from living things. ‘Never Alive’ means that nothing about the item is alive or the item is made of materials that never were alive. For example, items like plastic, rocks, or electronics. In these groups, you will see the word ‘ONCE ALIVE’ on the left, and ‘NEVER ALIVE’ on the right underneath each picture. If the item fits ONCE ALIVE, press the button under your index finger. If the item fits NEVER ALIVE, press the button under your middle finger. When you indicate that an item fits ONCE ALIVE, a green border will appear around the picture. When you indicate that an item fits NEVER ALIVE, a red border will appear around the picture.” Overall, participants were 85% accurate in correctly classifying the pictures as alive or never alive.

## Procedure

All study procedures took place at the Institute of Living/Hartford Hospital and were approved by the Institutional Review Board of Hartford Hospital. Informed consent was obtained from all individual participants included in the study. After consenting to

participate, participants completed the intake interview, consisting of the DIAMOND, CGI, and HRS-I. If eligible for the study, HD group participants returned to the clinic approximately one week later for the fMRI session, during which they completed the discarding, acquiring, and control tasks. They also completed the self-report measures during this visit. The HC group completed the tasks and questionnaires on the same day as the intake. The entire fMRI session took approximately 90 minutes. The order of tasks was counterbalanced across participants. Participants received \$50 for completing the fMRI session.

A subset of participants in the HD group ( $n = 42$ ) went on to complete treatment as part of the larger protocol. Treatment consisted of 16 weekly sessions of CBT conducted in a group format; treatment procedures included psychoeducation about HD, decision-making and problem-solving skills, emotion regulation skills (e.g., distress tolerance), cognitive restructuring of maladaptive hoarding-related beliefs, motivational enhancement, and relapse prevention (Tolin, Worden, Wootton, & Gilliam, 2017). These HD group participants completed the computer task in the MRI scanner again after treatment, within one week of the final group session.

## Statistical Analyses

First, we evaluated the construct validity of the discarding and acquiring tasks in relation to HD diagnosis. Independent samples *t*-tests and *Cohen's d* effect sizes were used to compare the HC and HD groups on item decisions (e.g., number of items discarded in the discarding task). For reaction time, we used 2 (group: HC vs. HD) x 2 (task: experimental [discarding/acquiring] vs. control) mixed analyses of variance (ANOVAs) with partial  $\eta^2$  effect sizes to examine main effects and interactions. Next, we evaluated the convergent validity of the tasks in relation to clinician-administered and self-report measures of HD severity using Pearson correlations. Finally, to assess treatment-related changes in the tasks, we used paired samples *t*-tests and *Cohen's d* effect sizes to compare pre- and post-treatment task measures among patients who completed CBT for HD.

## Results

### Sample Characteristics

Sample characteristics and symptom severity are presented in Table 1. The sample was predominantly female ( $n = 85$ , 79%), with an average age of 54.06 ( $SD = 8.19$ ) years. The HD and HC groups did not differ in terms of mean age, proportion of female sex, proportion of non-White race, or proportion of Hispanic/Latino ethnicity.

### Group Comparisons

Table 2 reports descriptive statistics, hypothesis testing results, and effect sizes for item decisions and reaction times in the HD and HC groups. For the discarding task, the HD group kept more items and discarded fewer items than did the HC group, with large effect sizes. For reaction time on this task, there was a main effect of group,  $F(1, 106) = 10.39$ ,  $p = .002$ ,  $\eta^2p = .09$  and task,  $F(1, 106) = 42.50$ ,  $p < .001$ ,  $\eta^2p = .29$ , with the HD group slower to make decisions across tasks than the HC group. There was also a group by task



interaction,  $F(1, 106) = 5.26, p = .024, \eta^2p = .05$ , such that the HD group showed comparable reaction times across tasks. By contrast, the HC group had greater differences in reaction times on the discarding task as compared to the control task (see Figure 1). The HD group had higher anxiety/fear and sadness/regret ratings following the discarding task than the HC group.

On the acquiring task, the HD group acquired marginally more items and left significantly fewer items than did the HC group, with moderate effect sizes (see Table 2). As expected, effect size estimates for group comparisons on item decisions were larger for the discarding task than the acquiring task (see Table 2). For reaction time on the acquiring task, there was a main effect of group,  $F(1, 106) = 11.28, p = .001, \eta^2p = .10$  and task,  $F(1, 106) = 57.46, p < .001, \eta^2p = .35$ , with the HD group slower to make decisions than the HC group. There was no group by task interaction,  $F(1, 106) = 1.80, p = .182, \eta^2p = .02$ . Again, anxiety/fear and sadness/regret ratings were higher in the HD group compared to the HC group.

### Correlations between Task Measures and Related Constructs

Correlations are presented in Table 3. Number of items saved and reaction times on the discarding task correlated positively with the difficulty discarding subscale of the SI-R and the fear of decision-making subscale of the FIS. Similarly, items acquired and reaction times on the acquiring task correlated positively with the acquiring subscale of the SI-R. Item decisions and reaction times on both tasks were negatively correlated with the positive attitudes subscale of the FIS, indicating that more positive attitudes towards decision-making were associated with fewer items kept and acquired as well as less time spent making decisions about task items. These results support the convergent validity of the discarding and acquiring tasks with established HD severity measures.

### Changes in Task Measures following Treatment

Across the discarding and acquiring tasks, reaction time and items kept and acquired (respectively) decreased from pre- to post-treatment with moderate to large effect sizes (see Table 4). On average, HD group patients kept approximately 9 items on the discarding task at post-treatment, in comparison to 13 items at pre-treatment. By comparison, the HC group kept an average of 10 items on the discarding task, indicating that after treatment, patients were keeping a similar amount of items as the HC group. HD group patients acquired an average of 7 items on the acquiring task at post-treatment in comparison to 9 items at pre-treatment; the HC group also acquired an average of 7 items on the acquiring task, again showing that after treatment patients behaved similar to HCs on the task. Reaction times at post-treatment also decreased. On the discarding task, the average reaction time at post-treatment was 1.89 seconds, whereas in the HC group the average was 1.16 seconds. Similarly, on the acquiring task, the post-treatment average was 1.94 in comparison to 1.96 in the HC group. These results support the sensitivity of the tasks in detecting treatment-related changes in acquiring, difficulty discarding, and reaction times.

## Discussion

The current study supports the validity of the discarding and acquiring tasks in measuring the core symptoms of HD. As predicted, the HD group acquired more and discarded less than the HC group, suggesting that the tasks detected symptoms of compulsive acquiring and difficulty discarding, respectively. Consistent with our pilot work (Tolin, Stevens, Nave, et al., 2012), the discarding task better discriminated between HD patients and HCs than the acquiring task, although both tasks showed differences between groups. These findings support the construct validity of the tasks in relation to HD diagnosis. Extending our earlier pilot work using this task (Tolin, Stevens, Nave, et al., 2012), we found that HD patients had slower reaction times than did HCs when making item decisions. Slower reaction times were also found for the control task, suggesting that HD patients have difficulties with decision-making under conditions of symptom provocation but also in more general situations. These results are consistent with executive functioning deficits that have been observed in HD patients in prior studies, such as difficulties with sustained attention, working memory, and cognitive flexibility (Ayers, Dozier, Wetherell, Twamley, & Schiehser, 2016; Tolin, Villavicencio, Umbach, & Kurtz, 2011). Interestingly, the interaction effect demonstrated that the HC group had greater differences in reaction times on the discarding task as compared to the control task, with quicker response times for the discarding task. It is likely that making discarding decisions, particularly for pictures of old, worn, or broken items, took very little time for the HC group, whereas having to categorize items as “alive” or “never alive” actually took time to think about prior to responding. For instance, it may take time to consider whether a paper item, which itself is not alive but is made up of natural products that were alive at one time, is “alive” or “never alive”.

The discarding and acquiring tasks also demonstrated convergent validity with clinician-administered and self-report measures of HD severity (i.e., the HRS-I and SI-R, respectively); item decisions and reaction times were positively correlated with these measures. It should be noted that some bivariate correlations were statistically significant but small in magnitude (e.g., between the acquiring subscale of the SI-R and items acquired during the acquiring task,  $r = 0.24$ ), suggesting that the task may not have captured participants' idiosyncratic triggers for acquiring. Alternatively, there may be important differences between self-reported acquiring and measures of acquiring behaviors in real time. These findings highlight the potential importance of behavioral measures to obtain a more comprehensive picture of symptom severity, as has been suggested by previous authors (Haynes et al., 2008). Similarly, there were some differences between the discarding and acquiring tasks in terms of associations with the HRS-I, the only clinician-administered measure we used in the present study. Although reaction time during the acquiring task was positively correlated with the HRS-I, items acquired during this task failed to show a significant association with HRS-I scores. By contrast, both item decisions and reaction time during the discarding task were significantly positively correlated with the HRS-I. It could be that the discarding task is a more valid and reliable measure of hoarding behaviors than is the acquiring task. These results may also reflect differences in the importance of difficulty discarding vs. acquiring to the HD diagnosis and core features. Indeed, although compulsive

acquisition is common among HD patients (Frost et al., 2009), acquiring is not included in the DSM-5 criteria for HD.

The tasks also showed sensitivity to detect treatment-related changes in acquiring, saving, and response times. At post-treatment, HD patients who had completed CBT behaved similarly on the behavioral tasks as compared to the HC sample. Based on these findings, it appears that the discarding and acquiring tasks may be used as treatment progress and outcome measures, although more research is needed to clarify whether the tasks reflect “real world” decisionmaking that occurs in patients’ homes with their own possessions.

There are a number of potentially important implications of this work. First, as stated previously, validating a behavioral measure of acquiring and discarding may overcome certain limitations of self-report measures of HD symptoms (e.g., inaccurate reporting, low insight into difficulties). Second, these behavioral tasks may provide clear targets for intervention that have the potential to improve current interventions. For instance, because the task assesses acquiring and discarding decisions as well as reaction time, it may provide critical information about the decision-making process that may inform treatment development and refinement. Given that CBT, the only empirically supported treatment for HD, is only moderately effective (average response rates range from 24 to 43%; Tolin, Frost, Steketee, & Muroff, 2015), efforts to improve existing interventions are critically needed. Third, this behavioral task may be a useful treatment progress and outcome measure that is more feasible than those that require significant clinician time and effort (e.g., home visits, sessions at a potential acquiring location).

The current study had several limitations. First, our sample was limited in terms of racial and ethnic diversity, leaving a gap in our understanding of how individuals from minority groups may respond to the discarding and acquiring tasks. Second, although we found positive associations between the discarding and acquiring tasks and established measures of HD severity, all laboratory tasks come with a question as to exactly how well they can predict “real world” behavior. Indeed, it is possible that HD patients showed improvements on the discarding and acquiring tasks after treatment, but had actually not improved on decision-making about personal possessions. It will be important to assess concordance between the task measures and actual discarding and acquiring decisions that occur in patients’ home environments. Establishing this concordance immediately after treatment and the ability of this behavioral paradigm to predict persistent treatment gains will be a valuable future research goal. Third, we did not assess inter-rater reliability for the DIAMOND, so we cannot be certain that the diagnoses assigned to participants were completely accurate. The DIAMOND has shown adequate inter-rater reliability for HD diagnoses and comorbid conditions in prior studies (Tolin et al., 2018).

With these limitations in mind, there are a number of potential future directions for this work. First, it will be important to replicate these findings in larger and more diverse samples, which will determine whether the tasks are reliable and valid for use with minority populations. Second, it will be interesting to examine the associations between task responses and actual acquiring and discarding decisions, particularly in the context of ongoing treatment. Should we find strong associations between the task and actual decision-

making, we will feel confident that the task accurately measures hoarding behaviors and treatment-related changes in these behaviors. For now, the results of the current study support the validity of the discarding and acquiring tasks in assessing HD behaviors. These tasks overcome certain limitations of existing HD measures by assessing discarding and acquiring decisions in real time, rather than via retrospective self-report. Other strengths of the current study include the use of both a clinical and comparison sample, a multimodal assessment approach comprised of clinician-administered measures, self-report measures, and behavioral measures, a longitudinal design, and an evidencebased treatment protocol.

## Acknowledgments

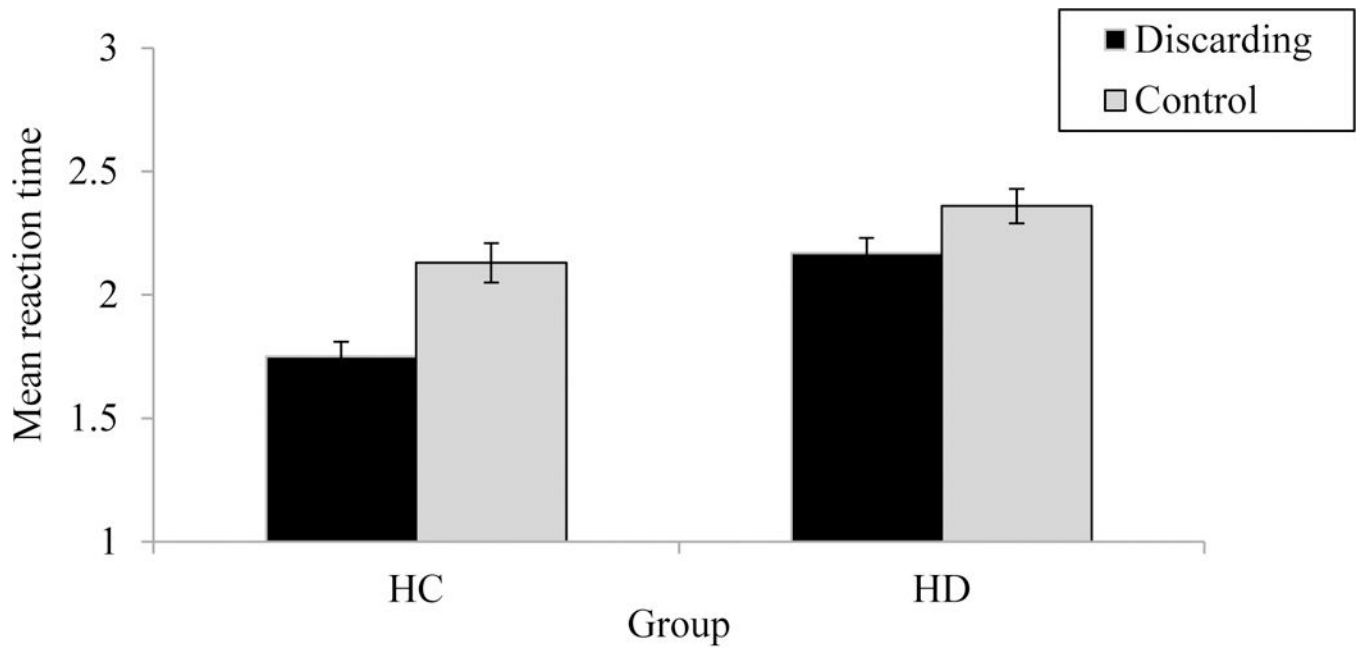
a. Funding: The study was funded by the National Institute of Mental Health (R01MH101163; PI: Tolin).

The authors wish to acknowledge Stephanie Preston for her assistance with the design of the computerized tasks.

## References

- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.
- Ayers CR, Dozier ME, Wetherell JL, Twamley EW, & Schiehser DM (2016). Executive functioning in participants over age of 50 with hoarding disorder. *Am J Geriatr Psychiatry*, 24(5), 342–349. doi: 10.1016/j.jagp.2015.10.009 [PubMed: 26809603]
- Ayers CR, Wetherell JL, Schiehser D, Almklov E, Golshan S, & Saxena S (2013). Executive functioning in older adults with hoarding disorder. *International Journal of Geriatric Psychiatry*, 28(11), 1175–1181. doi:10.1002/gps.3940 [PubMed: 23440720]
- Cogle JR, Wolitzky-Taylor KB, Lee HJ, & Telch MJ (2007). Mechanisms of change in ERP treatment of compulsive hand washing: does primary threat make a difference? *Behav Res Ther*, 45(7), 1449–1459. doi:10.1016/j.brat.2006.12.001 [PubMed: 17240352]
- DiMauro J, Tolin DF, Frost RO, & Steketee G (2013). Do people with hoarding disorder under-report their symptoms? *Journal of Obsessive-Compulsive and Related Disorders*, 2, 130–136. [PubMed: 23524977]
- Frost RO, & Shows DL (1993). The nature and measurement of compulsive indecisiveness. *Behav Res Ther*, 31(7), 683–692. [PubMed: 8216169]
- Frost RO, Steketee G, & Grisham J (2004). Measurement of compulsive hoarding: Saving Inventory-Revised. *Behaviour Research and Therapy*, 42(10), 1163–1182. [PubMed: 15350856]
- Frost RO, Tolin DF, Steketee G, Fitch KE, & Selbo-Bruns A (2009). Excessive acquisition in hoarding. *J Anxiety Disord*, 23(5), 632–639. [PubMed: 19261435]
- Frost RO, Tolin DF, Steketee G, & Oh M (2011). Indecisiveness and hoarding. *International Journal of Cognitive Therapy*, 4, 253–262.
- Grisham JR, Norberg MM, Williams AD, Certoma SP, & Kadib R (2010). Categorization and cognitive deficits in compulsive hoarding. *Behav Res Ther*, 48(9), 866–872. doi:10.1016/j.brat.2010.05.011 [PubMed: 20542489]
- Guy W (1976). *Assessment manual for psychopharmacology* Washington, DC: U.S. Government Printing Office.
- Haynes SN, Keawe'aimoku Kaholokula J, & Yoshioka DT (2008). Behavioral assessment. In Nezu AM & Maguth Nezu C (Eds.), *Evidence-based outcome research: A practical guide to conducting randomized controlled trials for psychosocial interventions* (pp. 6793). New York: Oxford University Press.
- Lawrence NS, Wooderson S, Mataix-Cols D, David R, Speckens A, & Phillips ML (2006). Decision making and set shifting impairments are associated with distinct symptom dimensions in obsessive-compulsive disorder. *Neuropsychology*, 20(4), 409–419. [PubMed: 16846259]

- Moshier SJ, Wootton BM, Bragdon LB, Tolin DF, Davis E, DiMauro J, & Diefenbach GJ (2016). The relationship between self-reported and objective neuropsychological impairments in patients with hoarding disorder. *Journal of Obsessive-Compulsive and Related Disorders*, 9, 9–15.
- Najmi S, Tobin AC, & Amir N (2012). Psychometric Properties of a Behavioral Test of Contamination-Related Obsessive-Compulsive Symptoms. *Cognitive Therapy and Research*, 36(3), 228–233. doi:10.1007/s10608-010-9332-0
- Preston SD, Muroff JR, & Wengrovitz SM (2009). Investigating the mechanisms of hoarding from an experimental perspective. *Depress Anxiety*, 26(5), 425–437. doi:10.1002/da.20417 [PubMed: 19242989]
- Rachman SJ (2004). Fear of contamination. *Behav Res Ther*, 42(11), 1227–1255. [PubMed: 15381436]
- Ruscio AM, Stein DJ, Chiu WT, & Kessler RC (2010). The epidemiology of obsessive-compulsive disorder in the National Comorbidity Survey Replication. *Mol Psychiatry*, 15(1), 53–63. doi: 10.1038/mp.2008.94 [PubMed: 18725912]
- Steketee G, & Frost RO (2003). Compulsive hoarding: Current status of the research. *Clinical Psychology Review*, 23, 905–927. [PubMed: 14624821]
- Tolin DF (2011). Understanding and treating hoarding: A biopsychosocial perspective. *J Clin Psychol*, 67(5), 517–526. doi:10.1002/jclp.20795 [PubMed: 21360705]
- Tolin DF, Frost RO, & Steketee G (2010). A brief interview for assessing compulsive hoarding: the Hoarding Rating Scale-Interview. *Psychiatry Research*, 178(1), 147–152. doi:10.1016/j.psychres.2009.05.001 [PubMed: 20452042]
- Tolin DF, Frost RO, Steketee G, & Muroff J (2015). Cognitive behavioral therapy for hoarding disorder: A meta-analysis. *Depression and Anxiety*, 32, 158–166. [PubMed: 25639467]
- Tolin DF, Gilliam C, Wootton BM, Bowe W, Bragdon LB, Davis E, ... Hallion LS (2018). Psychometric properties of a structured diagnostic interview for DSM-5 anxiety, mood, and obsessive-compulsive and related disorders. *Assessment*, 25(1), 3–13. doi: 10.1177/1073191116638410 [PubMed: 26988404]
- Tolin DF, Gilliam CM, Davis E, Springer K, Levy HC, Frost RO, & Steketee G (in press). Psychometric properties of the Hoarding Rating Scale-Interview. *Journal of Obsessive-Compulsive and Related Disorders*
- Tolin DF, Kiehl KA, Worchunsky P, Book GA, & Maltby N (2009). An exploratory study of the neural mechanisms of decision making in compulsive hoarding. *Psychol Med*, 39(2), 325–336. [PubMed: 18485263]
- Tolin DF, Stevens MC, Nave AM, Villavicencio A, & Morrison S (2012). Neural mechanisms of cognitive behavioral therapy response in hoarding disorder: A pilot study. *Journal of Obsessive-Compulsive and Related Disorders*, 1, 180–188.
- Tolin DF, Stevens MC, Villavicencio AL, Norberg MM, Calhoun VD, Frost RO, ... Pearson GD (2012). Neural mechanisms of decision making in hoarding disorder. *Arch Gen Psychiatry*, 69(8), 832–841. [PubMed: 22868937]
- Tolin DF, Villavicencio A, Umbach A, & Kurtz MM (2011). Neuropsychological functioning in hoarding disorder. *Psychiatry Research*, 189(3), 413–418. doi:10.1016/j.psychres.2011.06.022 [PubMed: 21764138]
- Tolin DF, Worden BL, Wootton BM, & Gilliam CM (2017). *CBT for hoarding disorder: A group therapy program: Therapist's guide* West Sussex, UK: Wiley-Blackwell.
- Worden BL, DiLoreto J, & Tolin DF (2014). Insight and motivation. In Frost RO & Steketee G (Eds.), *The Oxford handbook of hoarding and acquiring* New York: Oxford University Press.



**Figure 1.** Interaction between task and group on mean reaction time for the discarding task. Error bars represent standard errors. HC = Healthy control group. HD = Hoarding disorder group.



**Table 1**

## Demographic Characteristics and Symptom Severity

Variable	Full, <i>N</i> = 108		HD, <i>n</i> = 78		HC, <i>n</i> = 30		Comparison <i>t</i> or $X^2$ ( <i>p</i> )
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	
Age, <i>M</i> ( <i>SD</i> )	54.06	8.19	54.24	8.72	53.60	6.72	0.36 (.716)
Female sex	85	78.7	65	83.3	20	66.7	3.59 (.058)
Race							5.30 (.071)
White	97	89.8	73	93.6	24	80.0	
Black	8	7.4	3	3.8	5	16.7	
Asian	3	2.8	2	2.6	1	3.3	
Ethnicity							0.50 (.479)
Hispanic/Latino	2	1.9	1	1.3	1	3.3	
Not	106	98.1	77	98.7	29	96.7	
CGI, <i>M</i> ( <i>SD</i> )	3.96	2.01	5.15	0.89	1.07	0.25	4.08 (< .001)
HRS-I, <i>M</i> ( <i>SD</i> )	19.72	12.90	27.39	4.65	0.27	0.69	27.13 (< .001)

*Note.* HD = Hoarding disorder group. HC = Healthy control group. CGI = Clinical Global Impressions Scale. HRS-I = Hoarding Rate Scale – Interview. For CGI, mean comparison  $d = 6.24$ . For HRS-I, mean comparison  $d = 8.16$ .

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 2**

## Descriptive Statistics and Group Comparisons of Task Measures

Measure	HD Group, <i>M (SD)</i>	HC Group, <i>M (SD)</i>	<i>t(df)</i>	<i>p</i>	<i>d</i>
Discarding Task					
Items kept	13.19 (3.43)	10.67 (3.39)	3.43(106)	.001	0.74
Items discarded	16.42 (3.95)	19.23 (3.30)	-3.46(106)	.001	-0.77
Discarding task RT	2.17 (0.51)	1.75 (0.30)	5.32(86.94)	<.001	1.00
Anxiety/fear rating	21.17 (22.39)	1.17 (4.86)	7.45(93.29)	<.001	1.23
Sadness/regret rating	14.53 (18.23)	0.03 (0.18)	7.02(77.04)	<.001	1.12
Control task RT	2.36 (0.58)	2.13 (0.46)	1.86(106)	.066	0.44
Acquiring Task					
Items acquired	8.71 (3.51)	7.43 (3.41)	1.70(106)	.092	0.37
Items left	20.90 (3.90)	22.53 (3.48)	-2.01(106)	.047	-0.44
Acquiring task RT	2.38 (0.53)	1.96 (0.36)	3.97(106)	<.001	0.93
Anxiety/fear rating	16.41 (18.91)	2.00 (6.51)	5.88(105.24)	<.001	1.02
Sadness/regret rating	8.92 (17.84)	0.00 (0.00)	4.42(77.00)	<.001	0.71
Control task RT	2.65 (0.62)	2.35 (0.46)	2.40(106)	.018	0.55

*Note.* HD = Hoarding disorder. HC = Healthy control. RT = Reaction time. Discarding task RT and Acquiring task RT reflect average reaction times across keep and discard and acquired and left decisions, respectively. Positive *Cohen's d* values reflect greater numbers of items kept and acquired as well as greater reaction times in the HD group as compared to the HC group. Negative *Cohen's d* values reflect lower numbers of items discarded and left in the HD group as compared to the HC group.

**Table 3**

## Correlations between Task Measures and Hoarding Severity

Measure	Discarding Task		Acquiring Task	
	Items Kept	Task RT	Items Acquired	Task RT
<i>Clinician-Administered</i>				
HRS-I	0.33**	0.41***	0.18	0.38***
<i>Self-Report</i>				
SI-R Acquiring	0.37***	0.31**	0.24*	0.29**
SI-R Discarding	0.29**	0.41***	0.15	0.36***
SI-R Clutter	0.32**	0.39***	0.15	0.38***
SI-R Total	0.33**	0.39***	0.17	0.36***
FIS-Fear	0.26**	0.39***	0.26**	0.37***
FIS-Positive	-0.27**	-0.25*	-0.27**	-0.23*
<i>Visual Analogue Scale</i>				
Fear	0.17	0.35***	0.22*	0.27**
Sadness/Regret	0.19*	0.42***	0.35***	0.30**

*Note.* SI-R = Saving Inventory-Revised. HRS-I = Hoarding Rating Scale-Interview. FIS-Fear = Frost Indecisiveness Scale, Fear subscale. FIS-Positive = Frost Indecisiveness Scale, Positive attitudes subscale. RT = Reaction time. Discarding task RT and Acquiring task RT reflect average reaction times across keep and discard and acquired and left decisions, respectively.

\*  $p < .05$ .

\*\*  $p < .013$ .

\*\*\*  $p < .001$ .

**Table 4**

## Changes in Task Measures from Pre- to Post-Treatment

Measure	Pre, <i>M (SD)</i>	Post, <i>M (SD)</i>	<i>t(df)</i>	<i>p</i>	<i>d</i>
Discarding Task					
Items kept	12.93 (3.62)	8.93 (3.79)	5.08(40)	<.001	0.84
Discarding task RT	2.21 (0.57)	1.86 (0.43)	5.08(40)	<.001	0.79
Acquiring Task					
Items acquired	8.73 (3.11)	6.70 (3.98)	2.61(39)	.013	0.59
Acquiring task RT	2.40 (0.63)	1.93 (0.46)	4.80(39)	<.001	0.79

*Note.* RT = Reaction time.