Development and Validation of the Clutter Image Rating

**Randy O. Frost · Gail Steketee ·
David F. Tolin · Stefanie Renaud**

Abstract Few instruments are available to assess compulsive hoarding and severity of clutter. Accuracy of assessment is important to understanding the clinical significance of the problem. To overcome problems associated with over- and under-reporting of hoarding symptoms, the clutter image rating (CIR) was developed. This pictorial scale contains nine equidistant photographs of severity of clutter representing each of three main rooms of most people's homes: living room, kitchen, and bedroom. The psychometric properties of this measure were examined in two studies. Internal consistency, test–retest, and inter-observer reliabilities were good and convergent validity with other questionnaire and interview measures was also good. The CIR correlated more strongly with measures of clutter than with other hoarding and psychopathology scales. The CIR's very brief pictorial assessment method makes it useful in clinical and treatment contexts for measuring the clutter dimension of compulsive hoarding.

The Clutter Image Rating (CIR) is available in conjunction with the following publication: Steketee, G., & Frost, R.O. (2007). *Treatment of compulsive hoarding*. New York: Oxford University Press at www. oup.com.

R. O. Frost (✉) · S. Renaud
Department of Psychology, Smith College,
Northampton, MA 01063, USA
e-mail: rfrost@smith.edu

G. Steketee
Boston University,
Boston, MA, USA

D. F. Tolin
Institute of Living and University of Connecticut
School of Medicine,
Storrs, CT, USA

Keywords Hoarding · Collecting · Saving ·
Visual assessment · Testing

Compulsive hoarding consists of the acquisition of, and failure to discard, large numbers of material possessions resulting in clutter severe enough to cause emotional distress, impair functioning, and preclude the use of living spaces for their intended purposes (Frost and Hartl 1996; Frost et al. 2000a, b; Samuels et al. 2002, 2007). Hoarding has been reported in a variety of Axis I and II disorders including schizophrenia, social phobia, brain injury, depression, eating disorders, organic mental disorder, and dementia, as well as avoidant, dependent, and obsessive compulsive personality disorder (Steketee and Frost 2003). Hoarding behavior has been found in some forms of developmental disabilities as well, such as Prader–Willi syndrome (Dykens and Leckman 1996). Most often, however, it has been considered a symptom of obsessive compulsive disorder (OCD), occurring in 20–30% of OCD cases (Steketee and Frost 2003). However, several lines of evidence suggest that it is either a distinct subtype (e.g., McKay et al. 2004) or a separate disorder. A large percentage of people with hoarding problems experience no other OCD symptoms (Frost et al. 2006). Hoarding symptoms do not correlate as highly with other OCD symptoms as other symptoms do with each other (Wu and Watson 2005). Furthermore, hoarding may be as prevalent in patients with other anxiety disorders as it is in patients with OCD (Meunier et al. 2006). Hoarding patients show different patterns of cerebral glucose metabolism than do non-hoarding OCD patients (Saxena et al. 2004). Finally, treatments that work for other OCD symptoms (both medications and cognitive behavior therapy) appear less effective for hoarding (see Steketee and Frost 2003, for a review).

Hoarding can range from mild with little or no interference to life threatening, jeopardizing not only the health and safety of the sufferer, but those living nearby. Health department officials who have dealt with such cases reported that hoarding posed substantial health risks (Frost et al. 2000a, b). In 6% of the cases described by these officials, the hoarding contributed directly to the individuals' deaths in house fires. Examples of hoarding cases can be found in most communities and demonstrate the severity of this little understood syndrome. For instance, Jack was a 59-year-old former engineer who spent decades scavenging and buying sale items he intended to sell for a profit. Once in his home, however, he was unable to part with them. They overtook the house leaving him with literally no room to sit or sleep. His acquisitions also depleted his income, resulting in the shutting off of all his utilities (electricity, water, and gas). Eventually, the health department condemned his house as unfit for habitation. In another case, Jane, a 35-year-old single mother, was unable to discard anything that entered her home, even food containers, used band aids, or hair. The ensuing squalor led local child protective services staff to remove her two children until she got help.

Research on hoarding has been hampered by a lack of adequate measures. Many studies (e.g., Abramowitz et al. 2003) have relied on the Yale–Brown Obsessive-Compulsive Scale (Y-BOCS; Goodman et al. 1989) to assess hoarding. Unfortunately, the Y-BOCS symptom checklist contains only two yes/no items corresponding to hoarding obsessions and compulsions. These categorical judgments convey little information about the behavior, and the description given in the checklist does not mention cluttered living spaces as a symptom. Furthermore, the meaning of the Y-BOCS responses is distorted by aggregation with other OCD symptoms before making severity ratings, and the severity questions do not lend themselves well to accurate assessment of hoarding. Newer self-report scales for OCD (e.g., Vancouver Obsessive-Compulsive Inventory, Thordarson et al. 2004; Schedule of Compulsions, Obsessions, and Pathological Impulses, Watson and Wu 2005; Obsessive-Compulsive Inventory—Revised, Foa et al. 2002) improve on the Y-BOCS by including more items relevant to hoarding symptoms. However, they fail to capture the full range of the problem and have yet to be validated for hoarding (for example, on the Obsessive-Compulsive Inventory—Revised, the hoarding subscale was the only subscale on which healthy controls actually scored *higher* than did OCD patients).

The Saving Inventory—Revised (SI-R; Frost et al. 2004) is a self-report inventory that measures three primary components of hoarding—difficulty discarding, compulsive acquisition, and clutter. The SI-R contains 23 items which are scored for three subscales and a total score. Several recent studies indicate that the SI-R is reliable and can discriminate identified hoarding cases from non-hoarding controls and non-hoarding OCD cases (Frost et al. 2004). The subscales also correlate strongly with measures of hoarding-related beliefs and attachments, activity dysfunction resulting from hoarding behavior, and self and observer ratings of clutter in the home and are sensitive to treatment effects (Coles et al. 2003; Frost et al. 2004; Tolin et al. 2007). Several problems make development of additional measurement options beyond the current self-report inventories important. The literature on hoarding is replete with cases of patients' limited recognition of the problem (Christensen and Greist 2001; Damecour and Charron 1998; Fitzgerald 1997; Greenburg 1987; Shafran and Tallis 1996). Using various assessment methods, several studies reported that patients with hoarding symptoms showed less insight than did other non-hoarding patients (De Berardis et al. 2005; Frost et al. 1996; Samuels et al. 2007). Reports by public health officers and by elder service caseworkers indicated that fewer than 50% of hoarders recognized the severity of their problem (Frost et al. 2000a; Steketee et al. 2001). Many hoarding clients appear to ignore or not recognize the clutter in their homes, despite wading daily through a foot or more of debris throughout the home (Steketee and Frost 2003). Such clients often report not noticing the clutter when home alone, although they recognize it when the therapist points it out. To the extent that hoarding patients' reports are affected by this tendency to underestimate the severity of their hoarding behavior and its consequences, the validity of self-report inventories may be compromised.

Although poor insight may lead to underestimation of hoarding behavior, also problematic is the potential for *over*estimation of hoarding and clutter in particular. Several cases have been described in which individuals report severe hoarding behaviors and clutter in the home, but home visits failed to find significant clutter (Frost and Hartl 1996; Steketee et al. 2002). A similar pattern occurs in which some callers who identify themselves as "hoarders" prove to have relatively mild levels of clutter in their homes. Thus, variability in the use of the word "clutter" and in perceptions of clutter may make verbal self-reports of clutter inaccurate.

To resolve the problems associated with self-reports of the clutter dimension of hoarding behavior, a visual analogue to measure the severity of clutter in compulsive hoarding was developed and tested. This nine-point visual scale, the clutter image rating (CIR), consists of a series of nine photos of a room with increasing levels of clutter. Participants select the picture that best represents the clutter in the rooms of their own home. These pictorial representations require no descriptive language and avoid the problem of different definitions of clutter.

Study 1

Methods

Construction of the CIR

In order to generate the stimulus pictures, a small furnished apartment was rented and the bedroom, living room, and kitchen were filled with a wide variety of objects typically collected by people who hoard (newspapers, boxes, clothes, dishes, chairs, bottles, cans, books, pillows, televisions, cereal boxes, food containers, junk mail, etc.). These items were draped over piles of empty cardboard boxes to increase the volume of the materials, but only the hoarded materials were actually visible. The piles were built to within approximately 2 ft of the ceiling to mimic the most severe hoarding cases for each of the three rooms. Digital color photos were taken of each room with a wide-angle lens to fully capture the state of the room. After the first and each subsequent photo, several boxes beneath the surface were removed to reduce the volume of clutter while keeping materials on the surface as uniform as possible. This process continued until 30–35 pictures were taken for each room as the clutter was removed. The final picture contained no clutter.

Two procedures were employed to select the final set of nine photos for each room. First, the first author selected 22 of the 30–35 photos per room that appeared to represent the range of severity from uncluttered to extremely cluttered. Second, the least and most cluttered photos were used as bipolar anchors for the scaling task while the remaining photos were ordered from least to most cluttered. These 20 sequential photos were then grouped into five sequences of four pictures each. For example, pictures 1, 2, 3, and 4 formed the first group. One photo from each ordered grouping was randomly selected to create the first set of five photos. This procedure was repeated until all 20 pictures were assigned. The result was four sets of five randomly selected photos representing the full range of clutter levels. The original least and most cluttered photos were then affixed to each end of a table. The four sets of photos were used in a scaling task to select photos equally spaced in level of clutter.

Fourteen college students enrolled in an introductory psychology course participated in the scaling task. Participants were given the first set of five pictures, one after another, and asked to place each photo on the table at the appropriate distance relative to the least and most cluttered ones and to other photos in the set. After participants made final adjustments to the distances between photos, the experimenter recorded all distances and removed the photos. All participants repeated this sequence for each of the four sets of five pictures. The entire procedure was repeated for each of the three rooms.

The mean distance from the least cluttered photo was computed for each of the 20 photos. These means were graphed and used to select seven photos in addition to the least and most cluttered ones, so clutter distances were approximately equal between each of the nine photos. The resulting CIR scale included three nine-photo sequences with equal intervals between photos. Each set of photos depicting the living room, bedroom, and kitchen were arrayed on a laminated color page shown to research participants. Two studies were then conducted to examine the psychometric properties of the CIR.

Participants

Fifty-five people attending a workshop on clutter and hoarding were invited to participate in the study. Eight declined participation and one signed the consent form but did not complete any of the measures. The remaining 46 adults (33 women, 10 men, 3 failed to indicate gender) completed all study materials. Participants' ages ranged from 22 to 73 with a mean of 53.3 (SD=12.4). The majority were attending the workshop because of serious problems with hoarding and clutter. However, eight participants indicated they were family members or friends of people with hoarding problems. All participants completed the ratings based on their own behavior.

Measures

Clutter Image Rating (CIR) The CIR consisted of three pages of nine color photos representing a range of clutter in a living room (LR), bedroom (BR) and kitchen (K) as described earlier. Participants were asked to "select the picture that comes closest to the level of clutter in the corresponding room in your home." Scores ranged from 1 (least cluttered) to 9 (most cluttered). A mean composite score ranging from 1 to 9 was calculated across the three rooms for each person. Both composite and room-by-room scores were examined in this study.

Saving Inventory—Revised (SI-R) As noted earlier, this 23-item inventory (Frost et al. 2004) had three subscales: clutter (nine items), difficulty discarding and compulsive acquisition (7seven items each); items were scaled from 0 to 4. In the current sample the internal consistencies (Cronbach's α) ranged from 0.91 to 0.94.

Hoarding Rating Scale (HRS) Five questions from an interview (Tolin et al., submitted) protocol for obtaining clinician severity ratings were adapted to a questionnaire format. The questions asked about (1) difficulty using rooms in the home due to clutter, (2) difficulty discarding, (3) problems with collecting things, (4) the amount of

distress and (5) the amount of impairment these problems cause. These items were used individually in the analyses below.

Clutter Scale (CS) For each target room, participants were asked seven questions about the condition of the room (Hartl et al. 2002). The questions were: "To what extent: (1) does clutter in the room take up space intended for other purposes? (2) is the room neat? (3) are objects in the room efficiently organized? (4) would it be easy to find what one is looking for in the room? (5) is it difficult to walk through the room because of the clutter? (6) are the furniture tops cluttered? and (7) are the floor spaces cluttered?". Each question was scored on a seven-point scale with possible scores ranging from 7 to 49, higher scores reflecting more clutter. Internal consistencies for CS ratings of the 3 main rooms and total score were high (Kitchen $\alpha=0.91$; Living Room $\alpha=0.87$; Bedroom $\alpha=0.89$; Total $\alpha=0.94$).

Procedure

The present study was approved by the Institutional Review Board at Smith College. At the beginning of the workshop the attendees were invited to participate in a study of collecting, difficulty discarding, and clutter which would include completion of a set of questionnaires and ratings. Each participant received a packet of assessment materials including questionnaires and CIR pictures which were in a standard order. Those who wished to participate signed the consent form and completed the questionnaires at the beginning of the workshop.

Results

Mean scores for the SI-R-total (55.1, SD=19.2), clutter (23.6, SD=8.3), difficulty discarding (17.4, SD=6.7), and compulsive acquisition (14.1, SD=7.1) in this sample reflected moderately serious hoarding behavior compared to earlier samples (see Frost et al. 2004). CIR ratings ranged from 1 to 8 out of 9 for the bedroom (mean=3.49, SD= 1.7), from 1 to 7 for the living room (mean=3.21, SD=1.7), but only from 1 to 6 for the kitchen (mean=2.89, SD=1.2). Intercorrelations among the three CIR rooms were high (BR-LR $r=0.65$; BR-K $r=0.56$; LR-K $r=0.71$), and the internal consistency of the CIR composite score was acceptable ($\alpha=0.84$).

The CIR composite was highly correlated with both the SI-R clutter subscale ($r=0.72$) and HRS clutter ratings ($r=0.82$); further, the CIR was more weakly correlated with other subscales (r from 0.37 to 0.56). Rubin's Z tests (Meng et al. 1992) for the differences in the magnitude between these correlations indicated that the corresponding clutter correlations were significantly larger than CIR correlations with other subscales (all $p<0.05$).

Correlations between the CIR and Clutter Scale for corresponding rooms (r from 0.69 to 0.81) were larger than the correlations for non-corresponding rooms (r from 0.45 to 0.71). Six of the 12 comparisons were significantly larger ($p<0.05$) than for non-corresponding rooms (Rubin's Z tests for differences in dependent correlations) and were marginally larger for 2 others ($p<0.10$). Correlation of the CIR composite with the Clutter Scale total was very large ($r=0.79$).

Discussion

This initial study indicated that the CIR showed good internal consistency and good convergent and discriminant validity. It showed high correlations with other measures of clutter and weaker correlations with measures of related constructs including difficulty discarding, compulsive acquisition, and hoarding-related distress and interference. Further, correlations among the CIR room ratings and self-reported clutter in corresponding rooms were also high and generally larger than correlations with non-corresponding rooms.

The participants in this study were mostly people with hoarding problems, but a significant number (17%) did not have hoarding problems. The resulting distribution of scores covered the range from no hoarding to severe hoarding, but is heavily weighted toward people with hoarding problems.

Study 2

Although findings from Study 1 were promising, participants were not formally screened for hoarding symptoms and the measures against which the CIR was compared were all questionnaires and pertained only to hoarding behaviors. Accordingly, a second study was undertaken with participants who received a full diagnostic interview and completed various measures of hoarding, OCD, depression, and anxiety. Participants completed hoarding-related measures both in the clinic and at home where an experimenter also completed observational measures.

Methods

Participants

Adult participants (age 18 and over) were solicited through public service announcements offering a research/treatment opportunity. A trained interviewer administered the Anxiety Disorders Interview Schedule-Lifetime (ADIS-L; Brown et al. 1994), a structured interview to diagnose anxiety, mood,

and somatoform disorders. Items on hoarding (HRS) supplemented the standard ADIS interview. Participants were included if they received a severity rating of at least 4 ("definitely disturbing/disabling") on the clutter or difficulty discarding sections of the HRS hoarding ratings (see "Study 1" description). Participants were excluded for suicidal, psychotic or other symptoms requiring hospitalization. Participants were excluded if they presented evidence of mental retardation, dementia, brain damage, or severe cognitive dysfunction (score ≥8) on the Orientation-Memory-Concentration test (Katzman et al. 1983). No one was excluded due to these criteria.

Participants were 75 adults who qualified and completed the CIR in the clinic and/or at home. Of these, 39 participants were paid $20 per hour for their participation and 36 completed study measures as a condition of receiving free cognitive behavioral therapy for hoarding. The therapy was given as part of an open trial testing a newly developed treatment program for hoarding. The treatment project was underway when this project began, but ended before all the participants needed for this study were collected. Therefore, 39 people were paid for their participation rather than receiving treatment. Participants were diagnosed with a range of disorders, including 38 (51%) with major depressive disorder, 14 (19%) with non-hoarding OCD, 19 (25%) with generalized anxiety disorder, 23 (31%) with social phobia, and 9 (12%) with specific phobias. Only 6 participants (8%) failed to receive at least one diagnosis. Consistent with other studies of hoarding participants, most were Caucasian (91.4%), with 7.4% African American and 1.5% other. Women predominated (51; 68%), the mean age was 53.0 (SD=10.2, range 25–78), and only 44% were married or living with a partner. A large percentage of the participants completed high school (95.6%) and the majority completed college (58.5%). Only 58% of participants were employed at least part-time, and a surprisingly large percentage (15%) described themselves as disabled. Thirty-three percent of the participants reported incomes below $20,000, and 30% reported incomes above $50,000. The number of participants included in specific analyses varied as a consequence of missing data. Home visits were made for 58 participants. Participants declined home visits for a variety of reasons including distress over having someone see their home, unsuccessful attempts to arrange a time, and in one case, the house burned down shortly after the clinic assessment. There were no differences between sites on any of the demographic variables (age, gender, education, ethnicity; $p>0.05$).

Measures

Clutter Image Rating The CIR administration and scoring were as discussed for Study 1. In addition to the three main rooms (living room, bedroom, kitchen), additional rooms were rated by the participant in the clinic, the participant in their home, and the interviewer in the participant's home. Raters used the living room pictures as a proxy for assessing clutter in other locations (second bedroom/den, dining room, hallway, and car). As described in Study 1, a composite score was created by calculating the mean rating across the three rooms displayed in the photos. The internal consistency (α) for the participant in clinic was 0.80 ($n=69$), for participant in home it was 0.85 ($n=55$), and for therapist at home it was 0.89 ($n=56$).

Saving Inventory—Revised The SI-R (Frost et al. 2004), described previously, was administered in the clinic and at home for a subset of participants. In this sample the internal consistencies (α) for all three subscales and both contexts ranged from 0.80–0.89 ($n=36–70$).

Clutter Scale (CS) To simplify this measure, only three questions from the CS (Hartl et al. 2002) were examined for this study: clutter that takes up space intended for other purposes (no. 1) and occupies furniture tops (no. 6) and floor spaces (no. 7). These questions demonstrated the highest item-total correlations and had content more directly related to clutter. Participants completed the CS in the clinic and at home, and the interviewer did so in the participant's home. Internal consistencies (α) of ratings for the three main rooms (living room, bedroom, kitchen) were high: participant in clinic=0.89 ($n=66$), participant in home=0.88 ($n=43$), and therapist in home=0.93 ($n=46$). Ratings were also made of additional locations (second bedroom/den, dining room, hallway, and car).

Beck Depression Inventory—II (BDI–II) BDI-II (Beck et al. 1996) is a 21-item self-report measure of depressive symptoms with well-established reliability and validity. Participants completed the BDI-II at their initial session. The internal consistency (α) in the current study was 0.91.

Beck Anxiety Inventory (BAI) The BAI (Beck et al. 1988) is a 21-item self-report inventory of symptoms of anxiety with well-established reliability and validity. Participants completed the BAI at their initial session. The internal consistency (α) in the current study was 0.92.

Procedure

The present study was approved by the Institutional Review Boards at Hartford Hospital, Boston University, and Smith College. All participants signed a consent form before data collection began. Participants were screened by telephone and invited to come to the clinic for an initial diagnostic

assessment. Eligible participants also completed a battery of measures that day in the clinic. At a home appointment scheduled 1 week to 3 months later, participants and therapists/interviewers completed additional measures, including the CIR and SI-R.

Results

Scores on the CIR and SI-R

Means, standard deviations, and ranges for the CIR and SI-R can be found in Table 1, as well as correlations among the three CIR picture ratings in the clinic. The SI-R scores indicate substantial hoarding symptoms among these clinical participants (Frost et al. 2004), noticeably higher than in the Study 1 sample and spanning the full range of CIR scores. Mean scores for the participants indicated mild levels of depression and anxiety (see Table 1).

Reliability of the CIR

Retest reliability analyses for the participant CIR ratings completed in the clinic and at home for the 3 main rooms of the home can be found in Table 2. These were not purely test-retest correlations as they varied across context as well as time. Consequently they can be seen as a form of predictive validity as well. Correlations ranged from 0.62 to 0.81 for corresponding rooms, with an average of 0.73. The composite CIR showed a very high clinic/home correlation ($r=0.82$), which was similar to the correlation using only participants with a retest interval of 2 months or less ($r=0.85$).

Interobserver correlations of participant and experimenter ratings completed concurrently, but without consultation, at the participant's home can be found in Table 3. In addition to observer reliability, these correlations provide evidence of convergent validity as they compare participants' responses to a presumably objective standard (i.e., the experimenter). These correlations ranged from 0.73 to 0.94 for corresponding rooms and 0.94 for the composite measure.

Also in Table 3 are interobserver reliabilities for participant CIRs in the clinic and experimenter-rated CIRs at home. Although lower correlations would be expected because of the three sources of variance examined here (time, context, and observer), this information is important to understanding how closely participants' reports of clutter in the clinic are likely to match actual clutter in the home. These correlations ranged from 0.69 to 0.81 for corresponding rooms with an average of 0.75. The composite CIR showed good participant/experimenter correlation ($r=0.78$).

To compare the interobserver reliabilities of the CIR with those of the CS, comparable correlations were calculated for the CS total score. The correlation between the CS completed by the client and therapist at home was 0.77. This correlation was significantly smaller than the correlation between the client and therapist CIR completed in the home (0.94; $z=3.28$, $p<0.01$). Correlation of the CS completed by the client in the clinic and the therapist at home was 0.66, and did not differ statistically from the comparable CIR correlation (0.78, $z=1.2$, $p>0.05$).

Convergent Validity of the CIR

The Clutter Scale provided an alternative measure of clutter specific to each of the rooms measured by the CIR. The correlations between the Clutter Scale and CIR for each room are displayed in Table 4 for both the in-clinic assessment and in home assessment. The correlations for corresponding rooms ranged from 0.46 to 0.81. The mean

Table 1 Means, standard deviations, ranges, and inter-room correlations for the Clutter Image Rating (CIR), the Saving Inventory—Revised (SI-R), the Beck Depression Inventory (BDI), and the Beck Anxiety Inventory (BAI) in Study 2

	Mean	SD	Range	Bedroom	Living room	Kitchen
CIR						
Bedroom	4.34	2.16	1–9	–	–	
Living room	3.87	2.24	1–9	0.59*	–	–
Kitchen	3.79	2.01	1–9	0.41*	0.72*	
Composite	4.01	1.80	1.33–9	0.79*	0.91*	0.83*
SI-R						
Total	60.7	11.8	27–85			
Clutter	26.9	6.0	7–36			
Difficulty discarding	20.3	4.6	6–28			
Acquisition	15.5	5.8	3–26			
BDI	16.5	10.7	0–48			
BAI	10.4	9.5	0–51			

*$p<0.01$, $n=69$–73

Table 2 Correlations between participant clutter image ratings in the clinic and participant clutter image ratings at home for Study 2

Home CIR	Clinic CIR			
	Living room	Kitchen	Bedroom	Composite
Living Room	0.81	0.58	0.64	0.78
Kitchen	0.65	0.77	0.49	0.74
Bedroom	0.46	0.62	0.62	0.65
Composite	0.73	0.74	0.67	0.82

For all r, $p < 0.01$, 05, $n = 52$–54

correlation was 0.64, suggesting strong agreement between these room-specific measures of clutter. Correlations for home ratings tended to be slightly higher than those for clinic ratings, and living room correlations were strongest, followed by composite score correlations. The correlation between the composite measures (mean value of the CIR and Clutter Scale) was 0.77 for the home ratings and 0.70 for clinic ratings.

Table 5 shows similar correlations, but this time for the experimenter's CIR and Clutter Scale ratings, both conducted in the home. These correlations were very strong, ranging from 0.64 to 0.83 for the individual rooms and 0.78 for the composite measures.

CIR and CS correlations involving all seven rooms showed similar patterns of association, with higher correlations evident for matching rooms for both experimenter and participant. A seven-room composite score was highly correlated with the three-room composite (0.94) and revealed nearly identical correlations to the three-room composite for clinic versus home assessments (0.81 vs 0.82) and for participant versus experimenter (0.93 vs 0.94).

To further study convergent validity, the relationship of CIR composite ratings to the SI-R completed in the clinic (see Table 6) was also examined. As expected, the CIR in clinic and at home was more strongly correlated with participant ratings of the SI-R Clutter subscale ($r = 0.57$ to 0.63) than with the other SI-R subscales (Rubin's $Z > 2.49$; $p < 0.05$ for comparison of correlations).

Over/Underestimates of Clutter/Hoarding

While clinic-based hoarding measures are strongly correlated with home-based measures, differences in magnitude may occur in relation to the form of measurement (verbal self-report vs visual analogue; participant vs experimenter). To test these possibilities, hoarding measures administered in the clinic (CIR, Clutter Scale, SI-R) were compared to the same measures given in the home. Table 7 contains means, standard deviations, and t test values for these comparisons. The comparison for CIR home and clinic ratings was not significant ($p > 0.10$), but for the Clutter Scale, SI-R Clutter and SI-R Discarding subscales significantly higher scores occurred in the clinic compared to the home ($p < 0.04$). Differences in participant and experimenter ratings were not significant for the CIR, but on the Clutter Scale, participants rated their homes as more cluttered than did experimenters ($p < 0.03$).

CIR Relationship to Measures of Other Psychopathology

The CIR composite score was significantly but modestly correlated with the BAI ($r = 0.26$, $p < 0.05$), but not the BDI ($r = 0.13$, n.s.). Means for both measures reflected mild severity (BAI mean = 10.4, SD = 7.0; BDI mean = 16.5, SD = 10.7). Interestingly, the SI-R Clutter subscale was also uncorrelated with the BDI ($r = 0.21$, $p > 0.05$), while the Difficulty Discarding and Acquisition subscales were

Table 3 Correlations of clutter image ratings by participants (in the clinic and at home) with clutter image ratings by experimenters (in the home) for study 2

Experimenter CIR at home	Participant CIR							
	Living room		Kitchen		Bedroom		Composite	
	Clinic	Home	Clinic	Home	Clinic	Home	Clinic	Home
Living room	0.81	0.94	0.65	0.77	0.63	0.64	0.80	0.90
Kitchen	0.52	0.70	0.77	0.80	0.41	0.74	0.65	0.85
Bedroom	0.69	0.71	0.69	.60	0.69	0.73	0.71	0.79
Composite	.71	.88	.62	.80	.62	.77	**.78**	**.94**

All r significant at $p < 0.01$, $n = 50$–53

Table 4 Correlations of participants' clutter image ratings and clutter scale scores (in the clinic and at home) for study 2

Clutter scale	Clutter image rating							
	Living room		Kitchen		Bedroom		Composite	
	Clinic	Home	Clinic	Home	Clinic	Home	Clinic	Home
Living room								
Clinic	0.79		0.50		0.41		0.65	
Home		0.83		0.56		0.60		0.77
Kitchen								
Clinic	0.60		0.52		0.39		0.59	
Home		0.40*		0.78		0.47		0.61
Bedroom								
Clinic	0.35		0.35		0.46		0.45	
Home		0.42		0.33*		0.58		0.50
Composite								
Clinic	0.74		0.57		0.48		0.70	
Home		0.71		0.68		0.67		0.77

*$p<0.05$, otherwise all r significant at $p<0.01$, $n=59$–68 for clinic ratings and 39–46 for home ratings

correlated with it ($r=0.28$ and 0.28, $p<0.05$). All three SI-R subscales were correlated with the BAI, though the correlation with the SIR Acquisition subscale was higher ($r=0.52$, $p<0.01$) than with the other two ($r=0.28$ and 0.31, $p<0.05$).

Discussion

The present article describes the development and validation of a visual analogue measure of clutter for use in research and treatment of compulsive hoarding. This instrument showed good internal consistency and test-retest reliability, despite the fact that the retest varied across both context and time. Inter-observer reliability was excellent. The CIR demonstrated good convergent validity with other measures of clutter. It was highly correlated with both questionnaire and observer measures of room-by-room clutter, with the strongest relationships evident for corresponding rooms. Correlations with broader questionnaire and interview measures of hoarding revealed stronger

relationships with the clutter scores than with other dimensions of hoarding such as difficulty discarding and acquiring.

Although the CIR photos depict three rooms, using the living room pictures to assess other rooms in the home appeared to be successful. The three-room and seven-room composites were highly correlated and displayed a very similar pattern of correlations with other instruments. Thus, measuring clutter in more than the three main rooms depicted in the photos may not add much meaningful information.

Because few therapists in typical clinic settings can visit clients' homes, it is of special importance to understand the accuracy of clients' judgments made in the clinic of their clutter at home. In fact, participants' clinic and home CIR ratings were strongly correlated with each other, and with clinicians' ratings of the home at a later time. These findings suggest that the CIR completed in the clinic is a good representation of the clutter in the home. Although the CIR ratings in the home were done independently, it is possible that the close association between them was due to nonspecific effects of the experimenter being present when

Table 5 Correlations of experimenter ratings of clutter for Study 2 (both completed in the home)

Clutter scale	Clutter image rating			
	Living room	Kitchen	Bedroom	Composite
Living room	0.79	0.704	0.56	0.75
Kitchen	0.65	0.83	0.45	0.70
Bedroom	0.47	0.43	0.64	0.54
Composite	0.76	0.77	0.64	0.78

All r significant at $p<0.01$, $n=46$–51

Table 6 Correlations clutter image rating composites (in the clinic and at home) with the Saving Inventory—Revised

Measure	Participant/clinic CIR (n=67–71)	Participant /home CIR (n=54–57)	Experimenter CIR (n=52–56)
Saving Inventory—Revised			
Clutter	0.63**	0.57**	0.60**
Difficulty Discarding	0.28*	0.28*	0.29*
Acquisition	0.33**	0.29*	0.25

*p<0.05
**p<0.01

the CIR ratings were being done. However, the relatively high correlations between the experimenter ratings and participant ratings done in the clinic (despite the amount of time separating these ratings) argued against this possibility. Nonetheless, this may have influenced the magnitude of these correlations and the size of the discrepancy scores.

Hoarding clients may tend to overestimate clutter on questionnaire measures. On both the Clutter Scale and the SI-R clutter subscale, participants rated their clutter as significantly worse when they completed these measures in the clinic than when they did so at home. One possible explanation for this difference could be that participants cleaned or reduced the level of clutter from the first to the second administration of the measures. However, no such difference emerged in the CIR composite, suggesting that this rating bias stems from the nature of the measuring instrument (verbal versus visual). It also suggests that the CIR is not affected by over-reporting in the clinic, which is characteristic of paper and pencil measures.

The brevity of the CIR administration (less than 5 min) makes this reliable and valid measure a very useful screening tool to detect the presence of clinically significant hoarding symptoms. A cutoff score of 4 or higher can be used to indicate significant clutter requiring clinical attention. This measure may also be useful in assessing outcomes for interventions intended to reduce hoarding behavior. In an open trial of cognitive-behavior therapy for hoarding, the CIR was sensitive to treatment effects (Tolin et al. 2007).

Though the CIR holds promise as a measure of hoarding, in certain cases it can be misleading. For instance, occasionally people with hoarding problems live with or their homes are monitored by others (e.g., spouses, family members etc.) who prevent the buildup of clutter. In such cases the CIR, and other measures of clutter, would not accurately reflect the hoarding problem. Severity of clutter, as measured by the CIR, is only one dimension of hoarding. Severity of clutter may assess impairment of living spaces, but it is not a substitute for problems associated with difficulty discarding or excessive acquisition.

The absence of a significant correlation between the CIR and the BDI is interesting in light of the fact that over half of the sample received a major depressive disorder diagnosis. Previous studies have been inconsistent with respect to the relationship between hoarding and affect. While some have found small or insignificant association with negative affect or negative temperament (Grisham et al. 2005; Wu and Watson 2005), others have found significantly higher levels of depression among hoarding OCD patients compared to non-hoarding OCD patients (Frost et al. 2000b; Samuels et al. 2002). One reason for the inconsistency may be the feature of hoarding emphasized in each measure. Frost et al. (2004) found a different pattern of correlation with positive and negative affect for each SI-R subscale. Findings from the present study also demonstrated this inconsistency. Like the CIR, the SI-R clutter scale was not significantly correlated with the BDI while both the Difficulty Discarding and Acquisition subscales were. The Acquisition subscale also showed a stronger

Table 7 T tests across contexts for measures of hoarding and clutter for Study 2

Measure	Clinic	Home	Participant	Experimenter	t	df	p	d
Clutter image rating	4.05 (1.86)	3.81 (1.79)			1.60	53	0.12	0.22
Clutter scale	14.17 (3.72)	13.33 (4.14)			2.11	47	0.04	0.31
Saving Inventory								
Clutter	26.31 (4.53)	24.41 (6.25)			2.74	31	0.01	0.49
Discarding	19.56 (4.90)	18.19 (4.45)			2.49	31	0.02	0.44
Acquisition	14.81 (5.57)	14.94 (5.69)			0.21	31	0.84	0.04
Clutter image rating			3.80 (1.77)	3.70 (1.81)	1.13	52	0.26	0.15
Clutter scale			13.63 (4.12)	12.68 (4.32)	2.27	46	0.03	0.33

correlation with the BAI than did the other two subscales. The association of different dimensions of hoarding with negative affect deserves more research.

There are several limitations to the current study. The pictures used for the CIR were from relatively small rooms, they depicted commonly hoarded objects based on the authors' experience with hoarding clients, and the pictures displayed objects that were clean (i.e., not squalid conditions). These issues may limit the extent to which the CIR can be used for all cases of hoarding. Further research using the CIR is needed to determine applicability for these different contexts. Another potential limitation is that home visits were conducted for a subsample of the participants. It is possible that those who consented to a home visit were not representative of all participants. Finally, the clinic-home correlations cannot be considered pure test-retest reliabilities since they varied in context as well as time. A purer form of test-retest reliability would have been to complete the CIR twice in the same location.

References

Abramowitz, J. S., Franklin, M. E., Schwartz, S. A., & Furr, J. M. (2003). Symptom presentation and outcome of cognitive-behavioral therapy for obsessive-compulsive disorder. *Journal of Consulting and Clinical Psychology, 71*, 1049–1057.

Beck, A. T., Epstein, N., Brown, G., & Steer, R. A. (1988). An inventory for measuring clinical anxiety: Psychometric properties. *Journal of Consulting and Clinical Psychology, 56*, 893–897.

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Manual for the BDI—II.* San Antonio, TX: The Psychological Corporation.

Brown, T. A., DiNardo, P. A., & Barlow, D. H. (1994). *Anxiety disorders interview schedule for DSM-IV.* San Antonio, TX: The Psychological Corporation.

Christensen, D. D., & Greist, J. H. (2001). The challenge of obsessive-compulsive disorder hoarding. *Primary Psychiatry 8*, 79–86.

Coles, M. E., Frost, R. O., Heimberg, R. G., Steketee, G. (2003). Hoarding behaviors in a large college sample. *Behaviour Research and Therapy, 41*, 179–194.

Damecour, C. L., & Charron, M. (1998). Hoarding: A symptom, not a syndrome. *Journal of Clinical Psychiatry, 59*, 267–272.

De Berardis, D., Campanella, D., Gambi, F., Sepede, G., Salini, G., Carano, A. (2005). Insight and alexithymia in adult outpatients with obsessive-compulsive disorder. *European Archives of Psychiatry and Clinical Neuroscience, 255*, 350–358.

Dykens, E. M., & Leckman, J. F. (1996). Obsessions and compulsions in Prader–Willi syndrome. *Journal of Child Psychology and Psychiatry, 37*, 995–1002.

Fitzgerald, P. B. (1997). 'The bowerbird symptom': A case of severe hoarding of possessions. *Australian & New Zealand Journal of Psychiatry, 31*, 597–600.

Foa, E. B., Huppert, J. D., Leiberg, S., Langner, R., Kichic, R., Hajcak, G., et al. (2002). The Obsessive-Compulsive Inventory: Development and validation of a short version. *Psychological Assessment, 14*, 485–496.

Frost, R. O., & Hartl, T.L. (1996). A cognitive-behavioral model of compulsive hoarding. *Behavior Research and Therapy 14*, 341–350.

Frost, R. O., Krause, M. S., & Steketee, G. (1996). Hoarding and obsessive-compulsive symptoms. *Behavior Modification, 20*, 116–132.

Frost, R. O., Steketee, G., & Grisham, J. R. (2004). Measurement of compulsive hoarding, Saving Inventory—Revised. *Behaviour Research and Therapy, 42*, 1163–1182.

Frost, R. O., Steketee, G., Tolin, D., & Brown, T. (2006). *Co-morbidity and diagnostic issues in compulsive hoarding.* Paper presented at the Anxiety Disorders Association of America, Miami, FL.

Frost, R. O., Steketee, G., & Williams, L. (2000a). Hoarding: A community health problem. *Health and Social Care in the Community, 8*, 229–234.

Frost, R. O., Steketee, G., Williams, L., & Warren, R. (2000b). Mood, disability, and personality disorder symptoms in hoarding, obsessive compulsive disorder, and control subjects. *Behaviour Research and Therapy, 38*, 1071–1082.

Goodman, W. K., Price, L. H., Rasmussen, S. A., Mazure, C., Fleischmann, R. L., Hill, C. L., et al. (1989). The Yale–Brown Obsessive Compulsive Scale. I. Development, use, and reliability. *Archives of General Psychiatry, 46*, 1006–1011.

Greenburg, D. (1987). Compulsive hoarding. *American Journal of Psychotherapy, 41*, 409–416.

Grisham, J. R., Brown, T. A., Liverant, G. I., & Campbell-Sills, L. (2005). The distinctiveness of compulsive hoarding from obsessive-compulsive disorder. *Journal of Anxiety Disorders, 19*, 767–779.

Hartl, T. L., Allen, G. J., Frost, R. O., & Steketee, G. (2002). *Validation of a new measure of clutter to assess room and photographs.* Unpublished manuscript.

Katzman, R., Brown, T., Fuld, P., Peck, A., Schechter, R., & Schimmel, H. (1983). Validation of a short Orientation-Memory-Concentration Test of cognitive impairment. *American Journal of Psychiatry, 140*, 734–739.

McKay, D., Abramowitz, J. S., Calamari, J. E., Kyrios, M., Radomsky, A., Sookman, D., et al. (2004). A critical evaluation of obsessive-compulsive disorder subtypes: Symptoms versus mechanisms. *Clinical Psychology Review, 24*, 283–313.

Meng, X-L., Rosenthal, R., & Rubin, D. B. (1992). Comparing correlated correlation coefficients. *Psychological Bulletin, 111*, 172–175.

Meunier, S. A., Tolin, D. F., Frost, R. O., Steketee, G., & Brady, R. E. (2006). *Prevalence of hoarding symptoms across the anxiety disorders.* Paper presented at the Anxiety Disorders Association of America, Miami, FL.

Samuels, J., Bienvenu, O. J., III, Riddle, M. A., Cullen, B. A. M., Grados, M. A., Liang, K.-Y., et al. (2002). Hoarding in obsessive compulsive disorder: Results from a case-control study. *Behaviour Research and Therapy, 40*, 517–528.

Samuels, J., Bienvenu, O. J., III, Pinto, A., Fryer, A. J., McCracken, J. T., Rauch, S. L., et al. (2007). Hoarding in obsessive compulsive disorder: Results from the OCD Collaborative Genetics Study. *Behaviour Research and Therapy, 45*, 673–686.

Saxena, S., Brody, A. L., Maidment, K. M., Smith, E. C., Zohrabi, N., Katz, E., et al. (2004). Cerebral glucose metabolism in obsessive-compulsive hoarding. *American Journal of Psychiatry, 161*, 1038–1048.

Shafran, R., & Tallis, F. (1996). Obsessive-compulsive hoarding: A cognitive-behavioural approach. *Behavioural and Cognitive Psychotherapy, 24*, 209–221.

Steketee, G., de Noble, D., Diergerger, A., & Frost, R. O. (2002). *Compulsive hoarding among elderly people*. Unpublished manuscript.

Steketee, G., & Frost, R. O. (2003). Compulsive hoarding: Current status of the research. *Clinical Psychology Review, 23*, 905–927.

Steketee, G., Frost, R. O., & Kim, H-J. (2001). Hoarding by elderly people. *Health & Social Work, 26*, 176–184.

Thordarson, D. S., Radomsky, A. S., Rachman, S., Shafran, R., Sawchuk, C. N., & Hakstian, A. R. (2004). The Vancouver Obsessional Compulsive Inventory. *Behaviour Research and Therapy, 42*, 1289–1314.

Tolin, D. F., Frost, R.O., & Steketee, G. (2007). An open trial of cognitive-behavioral therapy for compulsive hoarding. *Behaviour Research and Therapy, 45*, 1461–1470.

Watson, D., & Wu, K. D. (2005). Development and Validation of the Schedule of Compulsions, Obsessions, and Pathological Impulses. *Assessment, 12*, 50–65.

Wu, K. D., & Watson, D. (2005). Hoarding and its relation to obsessive-compulsive disorder. *Behaviour Research and Therapy, 43*, 897–921.