University of Massachusetts Medical School

eScholarship@UMMS

Implementation Science and Practice Advances
Research Center Publications

Psychiatry

2007-1

Measuring homelessness and residential stability: The residential time-line follow-back inventory

Sam Tsemberis
Pathways to Housing

Et al.

Let us know how access to this document benefits you.

Follow this and additional works at: https://escholarship.umassmed.edu/psych_cmhsr

Part of the Health Services Research Commons, Psychiatric and Mental Health Commons, Psychiatry Commons, and the Psychiatry and Psychology Commons

MEASURING HOMELESSNESS AND RESIDENTIAL STABILITY: THE RESIDENTIAL TIME-LINE FOLLOW-BACK INVENTORY

Sam Tsemberis
Pathways to Housing

Gregory McHugo
New Hampshire-Dartmouth Psychiatric Research Center

Valerie Williams
Center for Mental Health Services Research, University of Massachusetts Medical School

Patricia Hanrahan
Department of Psychiatry, University of Chicago

Ana Stefancic
Department of Sociomedical Sciences, Columbia University

Reliable and valid longitudinal residential histories are needed to assess interventions to reduce homelessness and increase community tenure. This study examined the test-retest reliability, sensitivity to change, and concurrent validity of the Residential Time-Line Follow-Back (TLFB) Inventory, a method used to record residential histories in the Collaborative Program to Prevent Homelessness (n = 1,381). The Residential TLFB Inventory yielded temporally stable aggregate measures of duration in residential categories, and it revealed significant differences in change over time when contrasting study groups. A comparison of agency and participant data at one site

WILEY
InterScience®
DISCOVER SOMETHING GREAT

*demonstrated concurrent validity. These results support the psychometric
properties of the Residential TLFB Inventory and should encourage its use in
both clinical and research settings as a means to assess residential outcomes.*
© 2007 Wiley Periodicals, Inc.

Increasing rates of homelessness attest to the continual challenges faced by service providers (National Coalition for the Homeless, 2001; U.S. Conference of Mayors, 2001). Individuals with psychiatric disabilities—an estimated 20–25% of the single adult homeless population, 30–70% of whom have substance use disorders—present agencies with complex needs (Fischer & Breakey, 1991; Koegel, Sullivan, Burnham, Morton, & Wenzel, 1999; Rahav & Link, 1995). Due to the cyclic nature of long-term homelessness often experienced by this dually diagnosed population, longitudinal research is necessary to investigate patterns of homelessness and residential instability and the services that are best suited to address their needs (Hopper, Jost, Hay, Welber, & Haugland, 1997). To this end, it is necessary to obtain reliable and valid data regarding histories of residential stability and homelessness in order to assess the effectiveness of any intervention directed at reducing homelessness.

Studies on homelessness have differed, unfortunately, in their methods of measuring the key variables of homelessness and housing stability. Definitions of homelessness and housing stability have varied widely (Abdul Hamid, Wykes, & Stansfeld, 1993; Morrison, 1989), and the methodologies used to obtain data, as well as the samples studied, have been inconsistent (Newman, 2001). Studies of homelessness primarily use self-report and/or observer report to assess participants' homelessness or residential histories. Although several studies use institutional records or residential databases to supplement individual reports, such resources are often not readily available. The accuracy of retrospective self-report methods that extend over periods of several months may be hampered by challenges in one's ability to recall specifics over long time frames. Additional complications may include periods of intermittent homelessness in the interval between interviews, as well as frequent moves, substance use, and changes in psychological states (Carey, 1997; Klinkenberg et al., 2002). Nevertheless, with the right methodology, retrospective self-report, even among individuals with psychiatric disabilities or substance use disorders, has been shown to be reliable and valid (Calsyn, Allen, Morse, Smith, & Templehoff, 1993; Goldberg, Seybolt, & Lehman, 2002; Klinkenberg et al., 2002; Nieves, Draine, & Solomon, 2000; Sohler, Colson, Meyer-Bahlburg, & Susser, 2000).

Few studies describe the specific measurements used in collecting homelessness and residential information, thus limiting interpretations of validity and reliability (Goldfinger et al., 1997). The paucity of descriptions of instruments and interviewing methods has contributed to the lack of a standardized methodology for constructing participant residential chronologies and has seriously restricted our ability to compare outcomes across studies. This lack of standardized methodology compromises the usefulness and generalizabilty of research findings and may account for much of the variance observed in outcomes across studies.

Residential outcomes also have been measured in a variety of ways using different instruments and definitions. Two of the most frequently utilized variables have been residential status and residential stability. Residential status usually refers to the location at which participants were living at the time of follow-up, with studies usually reporting the percentage of participants living in a particular category at time of assessment. Assessments of residential stability provide a more complete evaluation of housing out-

comes over time. Some studies conceptualize stability as incorporating only the frequency of moves, whereas others take several factors into account, such as number of moves, frequency of moves, reasons for moves, type of residence (Bebout, Drake, Xie, McHugo, & Harris, 1997; Drake, Wallach, & Hoffman, 1989), and timing and entry into housing (Hurlburt, Hough, & Wood, 1996). Applying different combinations of these factors has led researchers to develop diverse categories of homelessness and residential stability (Brown, Ridgway, Anthony, & Rogers, 1991; Drake et al., 1989; Hurlburt et al., 1996).

Using uniform measures and examining similar housing outcomes would lead to better comparisons of residential stability across studies. Given the difficulties associated with self-report and the fact that few studies describe the specific methods used to elicit residential histories (Goldfinger et al., 1997), a measure is needed that utilizes a manageable time frame, provides optimal cues for aiding recall, and assesses many dimensions of housing outcomes.

THE RESIDENTIAL TIME-LINE FOLLOW-BACK INVENTORY

The Residential Time-Line Follow-Back (TLFB) Inventory (New Hampshire Dartmouth Psychiatric Research Center, 1995) presented here was used to record participants' housing histories across eight study sites in the Collaborative Program to Prevent Homelessness (CPPH). Designed to assess all dimensions of housing status and stability, the Residential TLFB Inventory incorporates both point-in-time assessments and longitudinal evaluations of housing and transitions to build a chronological record of each respondent's residential history for the period between successive interviews.

The TLFB method has its origins in the substance abuse field, where reliable and valid measures were sought for self-reported use of alcohol (Searles, Helzer, & Walter, 2000; Sobell, Maisto, Sobell, & Cooper, 1979; Sobell & Sobell, 1992; Sobell, Sobell, Klajner, Pavan, & Basian, 1986). The technique has continued to evolve, and it was used in the CPPH to assess alcohol and drug use, where its psychometric properties were examined (Sacks, Drake, Williams, Banks, & Herrell, 2003). The TLFB method also has been used in assessments of smoking behavior (Brown et al., 1998), binge eating (Bardone, Krahn, Goodman, & Searles, 2000), sexual behavior (Midanik et al., 1998; Weinhardt et al., 1998), and panic attacks (Nelson & Clum, 2002). Carey and colleagues (2001) demonstrated that the TLFB method is a reliable procedure among individuals with psychiatric disabilities.

The TLFB method was adapted for the assessment of residential history in order to capitalize on its strengths in improving recall from autobiographical memory. The Residential TLFB Inventory utilizes the TLFB interviewing method, including a calendar and memory cues, to elicit retrospective housing history. Reporting with accuracy where you have spent *each* night in the recent past (often 6 months) is a complex and error-prone cognitive task, especially if your recent residential history has been characterized by instability. Research has shown that memory for dates is poor and fades as time from the event passes (Bradburn, 2000). The TLFB method facilitates retrieval by utilizing Tourangeau's (2000) suggested methods for improving recall from autobiographical memory: 1) taking more time to remember, 2) decomposing a class of events into subclasses, 3) recalling events in reverse chronological order, and 4) listing boundaries or landmarks to aid recall.

This paper describes the Residential TLFB Inventory as applied to the collection of individual residential and homelessness data. It reports the Residential TLFB Inventory's test-retest reliability and concurrent validity for assessing housing outcomes, and it examines sensitivity to changes in residential status over time.

METHOD

Sites and Participants

The CPPH served adults with psychiatric and/or substance use disorders who were formerly homeless or at risk of becoming homeless (Rickards et al., 1999). Projects involved in the CPPH were located in the District of Columbia; Chicago, Illinois; Thornton, Colorado; Los Angeles, California; Tampa, Florida; New York City, New York; Philadelphia, Pennsylvania; and West Chester, Pennsylvania. Six of the eight projects primarily served individuals with psychiatric disabilities, and the other two served persons with substance use disorders. These projects provided and evaluated various combinations of housing programs, housing supports, case management, mental health treatment, substance abuse services, money management, and family interventions (Rickards et al., 1999).

The study populations across the sites varied in the extent of current and lifetime homelessness and the severity of mental illness and substance abuse. Half of the sites focused on preventing homelessness for those at risk, while the other half focused on reducing homelessness for people who were already homeless. The settings in which outreach and recruitment were conducted, and the enrollment procedures used, covered various points of entry into the service system from street outreach to community mental health agencies. Eligible participants were single adults aged 18 years or older who had a diagnosis of severe mental illness or substance use disorder, depending on the site's focus. Participants consisted of new program entrants who were referred from a variety of sources, many having histories of homelessness. All of the projects drew their treatment group participants from a single site. Four of the projects drew their comparison group participants from two sites, and the other four drew their comparison group participants from a single site. Three projects randomly assigned participants to treatment vs. comparison conditions, whereas the other five projects used a nonequivalent groups design.

The total sample included 1,381 participants at intake, 1,047 at 6 months, and 1,026 at 12-month follow-up. There were no significant differences in the demographic characteristics of the study group at these assessment points. Table 1 describes the study group at baseline. Overall, the measurement of homelessness was carried out with a richly varied array of sites, interventions, and samples.

Procedures and Measures

All interviewers were trained in general research interviewing techniques and on the study-specific instruments. All sites followed consent and confidentiality procedures developed by the CPPH Steering Committee, and all participants signed informed consent before taking part in the project. Participants were compensated for study assessments, with interview payments increasing over time from $15 at baseline to $25 at 12-month follow-up. There was an average of four interviewers per site, typically master's-level professionals. None of the sites used direct service provider staff as interviewers. The Residential TLFB Inventory was embedded in an omnibus cross-site instrument that took approximately 1.5 hours to administer and covered eight domains: participant demographics and history, physical health, housing stability and the extent of homelessness, substance use, mental health symptoms, perceived quality of life, service utilization and illegal activity, and contact with the legal system.

Table 1. Baseline Background Characteristics and Residential History (**n** = *1,381*)

	Total Cross-Site	Sample
Background Characteristics		
Mean age in years (SD)	40.3	(19.7)
% Female	42.2	
% Hispanic	10.1	
Race		
% African American	44.9	
% European American	39.8	
% Other	15.3	
Education		
% Less than HS diploma	37.8	
% HS diploma/GED	28.9	
% Some college or vocational/trade school	28.1	
% Four-year degree or higher	5.2	
Mean total Modified Colorado Symptom Index (MCSI) score at intake (SD)	20.8	(9.4)
% Used alcohol[a]	56.8	
% Used illicit drugs[a,b]	39.8	
Mean number of overnight psychiatric hospitalizations in lifetime (SD)	5.1	(12.7)
Residential History (Lifetime)[c]		
% Homeless at intake	52.0	
% Ever homeless	86.5	
Mean number of different times homeless (SD)	2.6	(1.4)
Mean years homeless in lifetime (SD)	3.5	(5.4)
Mean longest single period homeless in years (SD)	1.9	(3.7)

[a] Computed for the 6-month period prior to the baseline interview.
[b] Maximum number of days of use of one or more of the following drugs: sedatives, stimulants, opioids, cocaine, hallucinogens, and inhalants.
[c] Homelessness was defined as not having a regular place of residence and as living in a homeless shelter or living temporarily in an institution only because of having nowhere else to go, or living in a place not ordinarily used as a sleeping arrangement such as on the street, in a car, in an abandoned building, or in a bus or train station. This definition corresponds to the TLFB variable "literal homelessness."

Administration. The TLFB method begins with the interviewer noting events significant to the respondent, such as birthdays and holidays, on a 6-month calendar. The calendar serves as a visual aid to help respondents remember past living situations by placing them in time with the noted events, which serve as landmarks. Interviewers continue by inquiring where the participant lived during the past 6 months, beginning with the most recent living situation and working backward day by day to construct a residential history in reverse order. This entire exchange progresses without a prescribed script or time limitation. Landmark events and other prompts assist the respondent in placing each residential move in time: "Did you move into the shelter before or after your birthday? Was there snow on the ground when you entered the hospital?" Whereas dates may not be accessible in memory, durations may be, enabling the interviewer to develop a time line that is anchored by the interview date and any other dates that the respondent can provide. At the end of the reconstruction of the residential history, a forward run through the chronological sequence serves as a useful check on its validity. Finally, in the "Living Situation" section of the interview, the interviewer uses notes from the follow-back calendar to collect episodic information for each residence. Here, information is gathered for type and location of residence, date moved in and out, reasons for move, and household composi-

Table 2. Location Codes Used in the Cross-Site Instrument's Residential Location Grid

Item	Location Description	Final Categorization
1	All-night theater, subway station, or other public place	Literal Homelessness
2	Subway or bus	Literal Homelessness
3	Abandoned building	Literal Homelessness
4	Car or other private vehicle	Literal Homelessness
5	On the street or in other outdoor place	Literal Homelessness
6	Emergency shelter	Literal Homelessness
7	Hotel/motel	Temporary
8	Own SRO room (no services)	Stable
9	Someone else's SRO room (no services)	Temporary
10	Supportive SRO (services on site)	Stable
11	Drop-in center	Literal Homelessness
12	Safe haven (low-demand facility, reception center)	Temporary
13	Detox facility	Institutional
14	Crisis housing	Institutional
15	Intermediate care facility	Institutional
16	Own apartment or house	Stable
17	Parent/guardian's apartment or house (temporary)	Temporary
18	Parent/guardian's apartment or house (long-term)	Stable
19	Other family member's apartment or house (temporary)	Temporary
20	Other family member's apartment or house (long-term)	Stable
21	Someone else's apartment or house (temporary)	Temporary
22	Someone else's apartment or house (long-term)	Stable
23	Boarding house or board-and-care	Stable
24	Transitional housing program (short-term w/link to long-term)	Temporary
25	Transitional housing program (short-term w/o link to long-term)	Temporary
26	Transitional housing program (long-term)	Stable
27	Group home	Stable
28	Long-term alcohol/drug-free facility	Institutional
29	Hospital (medical only)	Institutional
30	Nursing home	Institutional
31	Treatment or recovery program	Institutional
32	Jail or prison	Institutional
33	Corrections halfway house	Temporary
34	Psychiatric hospital/facility (includes any inpatient psych. stay)	Institutional

tion. Although administration time can range from 5 to 45 minutes depending on the number of moves a participant has experienced, the average administration time for the Residential TLFB Inventory is approximately 15 minutes for a 6-month recall period.

All residential data are subsequently postcoded by the interviewer and recorded on the Residential Location Grid, allowing the information to be coded into a format that is more easily entered and analyzed. In the CPPH, each residence was coded using 1 of 34 location codes (see Table 2). For living arrangements that could be coded as either temporary or long-term, the distinction was made based on information elicited by the interviewer regarding the specific location. If participants indicated that they had intended to stay for a short time and had been living there for less than 6 months, the location was coded as temporary. If participants lived in a setting with no plans to move, or had lived there for more than 6 months, it was coded long-term.

Outcome Variables. The core CPPH cross-site residential outcome measures included literal homelessness, functional homelessness, and time in stable housing. To derive these three categories, the 34 location codes of the Residential TLFB Inventory were classified into one of the following mutually exclusive and exhaustive categories: 1) institutional residences, 2) stable residences, 3) temporary residences, and 4) literal homelessness (see Table 2).

To further refine the homelessness category, an additional definition of homelessness was developed—functional homelessness. For example, a participant's residential status was coded as functionally homeless for an entire period when she was literally homeless, institutionalized briefly, and then literally homeless again. Thus, functional homelessness refers to and counts time spent in settings classified as literally homeless as well as time in temporary and institutional settings, according to the type of residential locations that precede and follow the current setting. To systematize the classification of location codes as functionally homeless, the Residential subcommittee developed rules that detailed when a temporary and institutional setting would be considered functional homelessness. For example, a psychiatric hospitalization (institutional setting) would not be considered functionally homeless if the person was living in his or her own apartment (stable setting) prior to the hospitalization and returned there once discharged. However, if the participant was living on the street (literally homeless) prior to hospitalization (institutional setting) and returned to the street upon discharge, then she or he would be considered functionally homeless for that entire duration.

Test-Retest. At each site, interviewers invited early participants to take part in a second baseline interview until they recruited a total of 20 retest respondents. Interviewers re-interviewed those who agreed to participate in the retest 1–2 weeks after the baseline assessment (mean time between test and retest = 9.0 ± 5.0 days) using the entire cross-site instrument. The retest sample of 158 participants was recruited during the first 3 months of the study. Comparisons between the test-retest group ($n = 158$) and the remainder of the larger baseline group ($n = 1,223$) on four demographic variables and three homelessness variables revealed no significant differences at baseline. There were significant differences on two of the five substance use variables examined, days of alcohol use ($p < 0.05$) and number of drink equivalents ($p < 0.05$), with the test-retest group reporting less alcohol use.

Analysis

The cross-site Residential TLFB Inventory data were examined for reliability, sensitivity to change, and construct validity. Reliability estimates (kappas and intra-class correlations) and their confidence intervals were computed using procedures described by McGraw and Wong (1996). To check for differences in reliability by client characteristics, we generated reliability estimates for subgroups (e.g., men versus women) and used the confidence intervals to determine if differences were significant.

A nonparametric version of a paired *t*-test, the Wilcoxon signed-ranks test, was used to test for significant changes in the residential variables from baseline to 12 months. This test incorporates information about both the sign of the differences and the magnitude of the differences between pairs of values (Wilcoxon, 1945), without imposing restrictions on the distributional form of the variables. Comparisons between mean differences for the various residential measures by site type (prevention versus reduction) were made using the Mann-Whitney *U* test.

Table 3. Outcome Measures Computed from the Residential TLFB Calendar for the 6-Month Period Prior to the Baseline (n = 1,380)

	Mean (SD) at baseline						
Residential measure (days)[a]	*All sites* (n = 1,380)		*Prevention sites* (n = 597)		*Reduction sites* (n = 783)		p[b]
Literal homelessness[a]	42.4	(62.9)	10.8	(34.0)	66.4	(69.0)	< 0.001
Temporary settings[a]	26.2	(46.5)	18.3	(43.0)	32.3	(48.1)	< 0.001
Institutional settings[a]	36.7	(55.8)	28.5	(51.8)	42.9	(57.9)	< 0.001
Stable residences[a]	76.4	(78.1)	124.8	(71.3)	39.6	(61.0)	< 0.001
Functional homelessness	73.6	(75.3)	30.4	(54.9)	106.5	(72.1)	< 0.001
Time at current setting[c]	61.5	(70.8)	89.0	(76.2)	40.5	(58.3)	< 0.001
Total number of residences	5.0	(5.3)	3.0	(3.1)	6.6	(5.9)	< 0.001

[a] The literal, temporary, institutional, and stable categories sum to 100% of the duration. Participants with ≤ 90 days of the period accounted for in their grids were excluded from the analysis. Because of the way functional homelessness was defined, it will always be greater than or equal to the number of days literally homeless.
[b] Statistical significance of comparisons of prevention and reduction sites on each residential measure using the Mann-Whitney U test.
[c] Number of days reported at the location participant was living in at the time of the baseline interview.

Concurrent validity was determined for one site (Washington, D.C.), which contrasted the supported housing and continuum housing models in their effectiveness for persons with severe mental illness and current homelessness (McHugo et al., 2004). The agency using the continuum model controlled much of the housing offered to study participants, and thus kept records of occupancy and residential status for their clients throughout the study period. Agency data were obtained for 47 respondents for purposes of comparison to participant self-report data, which were collected using the Residential TLFB Inventory. Using the dates of the baseline and 6-month follow-up assessment, we were able to compare aggregate housing variables from the agency data with the self-report data.

RESULTS

Residential Stability at Baseline

Table 3 shows the average number of days spent in various categories of residence as reported by participants across all sites ($n = 1,380$) and the two site groupings, homelessness prevention and reduction, for the 6-month period prior to the baseline interview. At the cross-site level, participants reported the greatest number of days (76 days) in settings categorized as stable and the fewest (26 days) in those classified as temporary. This pattern varied, however, by site grouping. On average, respondents at prevention sites reported spending the majority of the 6-month period (125 days) prior to baseline in stable residences. Conversely, participants at reduction sites averaged only about 40 days in stable settings and spent the greatest number of days (66 days) literally homeless. Participants at reduction sites were also characterized by more days reported in temporary and institutional settings and more days functionally homeless, compared to those at prevention sites. At the cross-site level, participants had spent an average of 62 days in the setting where they reported living at time of baseline. They also reported an average of five different residences during the 6 months prior to baseline.

Table 4. Test-Retest Reliability (Intra-Class Correlation Coefficient, ICC) of Residential Measures Computed from the Residential TLFB Inventory for the 6-Month Period Prior to the Baseline (n = 157)[a]

Residential Measure (days)	ICC	(95% CI)
Literal homelessness	0.91	(0.88–0.94)
Temporary settings	0.59	(0.48–0.68)
Institutional settings	0.93	(0.90–0.95)
Stable residences	0.80	(0.74–0.85)
Functional homelessness	0.81	(0.75–0.86)
Current residence[b]	0.87	(0.83–0.90)
Most temporally distant residence[c]	0.59	(0.40–0.71)
Total number of residences	0.91	(0.87–0.93)

[a] One individual (Project H.O.M.E.) with responses that varied widely from test to retest was excluded from the analysis.
[b] Days reported at the location where the participant was living at the time of the baseline interview (first setting in grid).
[c] Days reported at the location where the participant was living at the time farthest from baseline interview (last setting in grid).

Test-Retest Reliability of Measures

Table 4 shows that the test-retest reliability coefficients were relatively high across the residential outcome measures (ICC = 0.80–0.93), with the exception of "days in temporary settings" (ICC = 0.59). The reliability of "days at current residence" (ICC = 0.87) was significantly higher than that of "days at most temporally distant residence" (ICC = 0.59). Test-retest reliability also varied by site grouping on two measures: for literal homelessness (prevention ICC = 0.44, 95% CI = 0.24–0.60; reduction ICC = 0.88, 95% CI = 0.82–0.92) and for time in institutional settings (prevention ICC = 0.95, 95% CI = 0.93–0.97; reduction ICC = 0.84, 95% CI = 0.76–0.89).

Test-retest reliability of the composite variables was unrelated to age, gender, race/ethnicity, level of psychiatric symptomatology at intake (measured by a modified version of the Colorado Symptom Index), lifetime homelessness in months, or total number of residences during the 6 months prior to the baseline interview. For the "days in institutional settings" variable, however, better reliability was associated with fewer days of reported alcohol use ($p < 0.001$), fewer days of reported cocaine use ($p < 0.001$), fewer days of reported cannabis use ($p < 0.02$), and more reported psychiatric hospitalizations ($p < 0.01$).

Sensitivity to Change

Sensitivity to change was assessed by comparing the residential outcome measures obtained for the 6 months prior to the baseline interview to the same measures for the 6 months prior to the 12-month follow-up interview. Due to attrition at 12-month follow-up, these analyses are based on a reduced study group. There were no differences in baseline demographic characteristics between the study group and the dropouts at 12 months, but there were significant differences at baseline for days spent literally homeless ($p < 0.005$), functionally homeless ($p < 0.02$), and in stable settings ($p < 0.02$). Participants who were lost to follow-up at 12 months had spent more days homeless and fewer days stably housed when they entered the study.

These analyses were based on the proportion of days in a given residential setting, rather than the raw number of days. Although participants' baseline Residential TLFB Inventory was based on 6 months, their residential data at 12-month follow-up was based

on a variable number of days since their prior interview. Because few participants were interviewed on the ideal follow-up date, and because some had missed the 6-month interview, the recall period at 12-month follow-up varied from 91 days to 368 days (mean = 202.9 days; median = 183 days; SD = 48.2). To deal with this large range, the residential outcome measures were converted to proportions for each participant, by dividing the number of days in a given setting by the number of days in the recall period. Changes over time were significant for the total cross-site sample ($n = 987$) on all of the residential measures (see Table 5). Participants at reduction sites generally exhibited greater gains from baseline to one year than those at prevention sites. For the entire cross-site sample, as well as for the grouped sites, the biggest changes were reduction in the proportion of time spent functionally homeless and increases in the proportion of time in stable residences.

Validity

The Washington, D.C. site compared the average number of days reported by the participant ($n = 47$) and the agency for literal homelessness, stable housing, temporary settings, and institutional settings during the first 6 months of the study. The means within each residential category were similar across the two sources and the Pearson r's between agency and participant data were high, ranging from 0.84 for stable housing to 0.92 for literal homelessness. These results support the validity of the instrument when the residential data from specific settings are combined into aggregate categories over a 6-month period.

DISCUSSION

The Residential TLFB Inventory proved to be a valid and reliable method for measuring the key outcome variables in this large-scale multi-site study of homelessness and residential stability among individuals with psychiatric disabilities and/or substance use disorders. With the capacity to assess patterns of change over extended time frames, this tool offers a significant advance in the assessment of homelessness and residential stability, especially with its ability to capture individual variability in homelessness and residential patterns.

The convenience and economic efficiency of the retrospective method used by the Residential TLFB Inventory provides a considerable methodological advantage across a variety of clinical settings, and it provides the flexibility to compare homelessness and residential data across many sites at various time intervals. The calendar, history form, codes, and probes provide the structure necessary for standardizing the methodology, while the absence of a specific script allows for a wide range of applicability, including the ability to administer the measure to both providers and participants for purposes of making comparisons of residential outcomes. The statistically significant changes observed in the expected direction at both the cross-site and site-grouping levels indicated that the composite measures derived from the Residential TLFB Inventory were sensitive to change in homelessness and residential stability over time. Additionally, this tool affords the opportunity for further evaluating residential stability by examining number of moves, length of intervals between moves, reasons for moving, and residential composition. By gathering continuous day-by-day residential information, investigators can also use the cost associated with each particular residence to calculate the residential costs incurred by participants over time (Gulcur, Stefancic, Shinn, Tsemberis, & Fischer, 2003).

Table 5. Change in Residential Measures from Baseline to 12-Month Interview, Expressed as the Difference in the Proportion of Time in Five Aggregate Housing Categories

Residential Measure (proportion of 6-month period)[a]	Site Type			
	All sites (n = 987) Mean difference [b]	Prevention sites (n = 396) Mean difference [b]	Reduction sites (n = 591) Mean difference [b]	p [c]
Literal homelessness	−0.11[d]	−0.03[d]	−0.19[d]	<0.001
Temporary settings	−0.03[e]	−0.03[e]	−0.03[e]	0.990
Institutional settings	−0.05[d]	−0.03[e]	−0.07[d]	0.068
Stable residences	+0.15[d]	+0.08[d]	+0.22[d]	<0.001
Functional homelessness	−0.15[d]	−0.07[d]	−0.23[d]	<0.001

[a] This proportion represents the days in each category divided by the days in the total period. Participants with ≤ 90 days of the period accounted for were excluded, leading to the difference in sample size (1,026 vs. 987).
[b] Difference between the means of the measure at 12 months and at baseline.
[c] Comparison of prevention and reduction site mean differences.
[d] Wilcoxon signed-ranks asymptotic Z, $p < 0.001$.
[e] Wilcoxon signed-ranks asymptotic Z, $p < 0.05$.

The Residential TLFB Inventory allows for locally meaningful housing categories and assesses residential patterns over time. It is straightforward to create independent aggregate categories of residential settings, such as the four mutually exclusive categories used in the CPPH. It is also possible to create hybrid categories that include both residential and homelessness outcomes, such as functional homelessness, which resolves the ambiguity of classifying days spent in institutions or in temporary settings. Pursuant to the definition of functional homelessness, such ambiguous stays would be categorized in relation to the participant's preceding and subsequent living situations.

The test-retest reliability of the computed residential measures indicates that study participants understood the TLFB interview and were able and willing to respond consistently. There were two noteworthy qualifications. First, comparison of the test-retest reliability for the duration at the current residence vs. the most temporally distant residence reveals a drop in reliability as time between occupancy and recall increases. This finding is consistent with numerous studies of autobiographical memory, but its impact is reduced by using composite variables as outcome measures. Second, among the five aggregate residential outcomes used in CPPH, the number of days in temporary settings was less stable temporally than the others (0.59 vs. 0.80–0.93). Stays in temporary settings may be less memorable for participants, and it is likely that there was greater confusion among interviewers concerning the use of these location codes. Indeed, coding of individual residences into 1 of the 34 location categories was somewhat inconsistent across time and interviewers in the CPPH, particularly for transitional housing or other temporary congregate housing settings. On the one hand, the ambiguity of some of the residential categories led to these discrepancies, as interviewers could not conclusively determine whether a particular residence was short-term transitional, long-term transitional, group residence, and so forth. On the other hand, there were some important residential categories, such as living doubled up, that were not included explicitly among the 34 codes and had to be inferred from stays in temporary settings.

One way to remedy this problem would be to provide interviewers with a list of residential programs within the study area, along with their corresponding residential code,

thereby increasing test-retest reliability for the composite measure. Anticipating that there will be some residences not yet included on the list or that study participants will be found in settings not anticipated by the existing codes, the team should develop a protocol for determining the codes in cases where there is uncertainty about their categorization. When administering the instrument in particular locations, location codes can also be tailored to create fewer and more specific categories.

The high agreement found between one site's agency residential records and participant self-reports supports the validity of the composite variables that were calculated for CPPH from the Residential TLFB Inventory. Because this finding was based on only one group at one site, it is essential that further studies using this instrument include data from multiple sources in order to study validity further.

One final difficulty with the Residential TLFB Inventory was associated with attempts to standardize the follow-up time periods across participants. Due to difficulties associated with locating and scheduling participants within specified time frames, or to participants' missing entire interview cycles, not all interviews covered exactly 6 months. To address this problem, the calendar should be designed to account for as many months as is necessary to work backwards until the date of the previous interview. To compensate for the problem of varying recall periods, we converted the raw number of days in various settings or in aggregate categories of settings into proportions by dividing the number of days spent in each setting (or setting category) by the total number of days in the recall period. This procedure equates the outcome variables from individuals with different lengths of time between successive interviews. Another strategy for longitudinal studies is to use mixed-effects models and to treat time as a continuous, rather than a classification, variable. For example, each participant's follow-up points would be indicated by the number of weeks since baseline, and time would be specified as a continuous variable in the design. This approach is akin to treating time between interviews as a covariate in a repeated measures design (e.g., Toro et al., 1997).

Overall, this study provides support for the reliability and validity of the Residential TLFB Inventory as a means of assessing homelessness and residential stability among a variety of populations and settings. The results support the utility of this tool for homelessness research and housing studies, by providing a methodology for understanding various aspects of the longitudinal course of housing and homelessness.

REFERENCES

Abdul Hamid, W., Wykes, T., & Stansfeld, S. (1993). The homeless mentally ill: myths and realities. International Journal of Social Psychiatry, 39(4), 237–254.

Bardone, A.M., Krahn, D.D., Goodman, B.M., & Searles, J.S. (2000). Using interactive voice response technology and timeline follow-back methodology in studying binge eating and drinking behavior: Different answers to different forms of the same question? Addictive Behaviors, 25(1), 1–11.

Bebout, R.R., Drake, R.E., Xie, H., McHugo, G.J., & Harris, M. (1997). Housing status among formerly homeless, dually diagnosed adults in Washington, D.C. Psychiatric Services, 48, 936–941.

Bradburn, N.M. (2000). Temporal representation and event dating. In A.A. Stone, J.S. Turkkan, C.A. Bachrach, J.B. Jobe, H.S. Kurtzman, & V.S. Cain (Eds.), The Science of Self-Report: Implications for Research and Practice (pp. 49–61). Mahwah, NJ: Lawrence Erlbaum.

Brown, M.A., Ridgway, P., Anthony, W.A., & Rogers, E.S. (1991). Comparison of outcomes for clients seeking and assigned to supported housing services. Hospital & Community Psychiatry, 42(11), 1150–1153.

Brown, R.A., Burgess, E.S., Sales, S.D., Whiteley, J.A., Evans, D.M., & Miller, I.W. (1998). Reliability and validity of a smoking timeline follow-back interview. Psychology of Addictive Behaviors, 12(2), 101–112.

Calsyn, R.J., Allen, G., Morse, G.A., Smith, R., & Templehoff, B. (1993). Can you trust self-report data provided by homeless mentally ill individuals? Evaluation Review, 17(3), 353–366.

Carey, K.B. (1997). Reliability and validity of the time-line follow-back interview among psychiatric outpatients: A preliminary report. Psychology of Addictive Behaviors, 11(1), 26–33.

Carey, M.P., Carey, K.B., Maisto, S.A., Gordon, C.M., & Weinhardt, L.S. (2001). Assessing sexual risk behaviour with the Timeline Followback (TLFB) approach: Continued development and psychometric evaluation with psychiatric outpatients. International Journal of STD & AIDS, 12(6), 365–375.

Drake, R.E., Wallach, M.A., & Hoffman, J.S. (1989). Housing instability and homelessness among aftercare patients of an urban state hospital. Hospital & Community Psychiatry, 40(1): 46–51.

Fischer, P.J., & Breakey, W.R. (1991). The epidemiology of alcohol, drug, and mental disorders among homeless persons. American Psychologist, 46, 1115–1128.

Goldberg, R.W., Seybolt, D.C., & Lehman, A. (2002). Reliable self-report of health service use by individuals with serious mental illness. Psychiatric Services, 53(7), 879–881.

Goldfinger, S.M., Schutt, R.K., Tolomiczenko, G.S., Turner, W., Ware, N., Penk, W.E., et al. (1997). Housing persons who are homeless and mentally ill: Independent living or evolving consumer households? In W.R. Breakey & J.W. Thompson (Eds.), Mentally Ill and Homeless: Special Programs for Special Needs (pp. 29–49). Amsterdam: Harwood Academic Publishers.

Gulcur, L., Stefancic, A., Shinn, M., Tsemberis, S., & Fischer, S. (2003). Housing, hospitalization, and cost outcomes for homeless individuals with psychiatric disabilities participating in continuum of care and housing first programs. Journal of Community and Applied Social Psychology, 13(2), 171–186.

Hopper, K., Jost, J., Hay, T., Welber, S., & Haugland, G. (1997). Homelessness, severe mental illness, and the institutional circuit. Psychiatric Services, 48, 659–665.

Hurlburt, M.S., Hough, R.L., & Wood, P.A. (1996). Effects of substance abuse on housing stability of homeless mentally ill persons in supported housing. Psychiatric Services, 47, 731–736.

Klinkenberg, W.D., Calsyn, R.J., Morse, G.A., McCudden, S., Richmond, T.L., & Burger, G.K. (2002). Consistency of recall of sexual and drug-using behaviors for homeless persons with dual diagnosis. AIDS and Behavior, 6(4), 295–307.

Koegel, P., Sullivan, G., Burnham, A., Morton, S.C., & Wenzel, S. (1999). Utilization of mental health services and substance abuse services among homeless adults in Los Angeles. Medical Care, 27, 306–317.

McGraw, K.O., & Wong, S.P. (1996). Forming inferences about some intraclass correlation coefficients. Psychological Methods, 1(1), 30–46.

McHugo, G.J., Bebout, R.R., Harris, M., Cleghorn, S., Herring, G., Xie, H., et al. (2004). A randomized controlled trial of integrated versus parallel housing services for homeless adults with severe mental illness. Schizophrenia Bulletin, 30, 969–982.

Midanik, L.T., Hines, A.M., Barrett, D.C., Paul, J.P., Crosby, G.M., & Stall, R.D. (1998). Self-reports of alcohol use, drug use and sexual behavior: Expanding the Timeline Follow-back technique. Journal of Studies on Alcohol, 59(6), 681–689.

Morrison, J. (1989). Correlations between definitions of the homeless mentally ill population. Hospital & Community Psychiatry, 40(9), 952–954.

National Coalition for the Homeless. (2001). Bringing America Home: The Strategic Plan of the National Coalition for the Homeless. Washington, DC: National Coalition for the Homeless.

Nelson, W.A., & Clum, G.A. (2002). Assessment of panic frequency: Reliability and validity of a time-line follow-back method. Journal of Psychopathological Behavior, 24(1), 47–54.

New Hampshire Dartmouth Psychiatric Research Center. (June 1995). Residential Follow-Back Calendar. Lebanon, NH: Dartmouth Medical School.

Newman, S.J. (2001). Housing attributes and serious mental illness: Implications for research and practice. Psychiatric Services, 52(10), 1309–1317.

Nieves, K., Draine, J., & Solomon, P. (2000). The validity of self-reported criminal arrest history among clients of a psychiatric probation and parole service. Journal of Offender Rehabilitation, 30(3–4), 133–151.

Rahav, M., & Link, B.G. (1995). When social problems converge: Homelessness, mentally ill, chemical abusing men in New York City. International Journal of Addiction, 30, 1019–1042.

Rickards, L.D., Leginski, W., Randolph, F.L., Oakley, D., Herrell, J.M., & Gallagher, C. (1999). Cooperative agreements for CMHS/CSAT Collaborative Program to Prevent Homelessness: An overview. Alcoholism Treatment Quarterly, 17(1/2), 1–15.

Sacks, J.Y., Drake, R.E., Williams, V.F., Banks, S.M., & Herrell, J. (2003). Utility of the Time-Line Follow-Back to assess substance use among homeless adults. Journal of Nervous and Mental Disease, 191(3), 145–153.

Searles, J.S., Helzer, J.E., & Walter, D.E. (2000). Comparison of drinking patterns measured by daily reports and timeline follow back. Psychology of Addictive Behaviors, 14(3), 277–286.

Sobell, L.C., Maisto, S.A., Sobell, M.B., & Cooper, A.M. (1979). Reliability of alcohol abusers' self reports of drinking behavior. Behavior Research and Therapy, 17, 157–160.

Sobell, L.C., & Sobell, M.B. (1992). Timeline follow-back: A technique for assessing self-reported alcohol consumption. In R. Litten & J. Allen (Eds.), Measuring Alcohol Consumption. (pp. 41–72). Totowa, NJ: The Humana Press, Inc.

Sobell, M.B., Sobell, L.C., Klajner, F., Pavan, D., & Basian, E. (1986). The reliability of a timeline method for assessing normal drinker college students' recent drinking history: Utility for alcohol research. Addictive Behaviors, 11, 149–161.

Sohler, N., Colson, P.W., Meyer-Bahlburg, H.F., & Susser, E. (2000). Reliability of self-reports about sexual risk behavior for HIV among homeless men with severe mental illness. Psychiatric Services, 51(6), 814–816.

Toro, P.A., Rabideau, J.M.P., Bellavia, C.W., Daeschler, C.V., Wall, D.D., Thomas, D.M., et al. (1997). Evaluating an intervention for homeless persons: Results of a field experiment. Journal of Consulting and Clinical Psychology, 65, 476–484.

Tourangeau, R. (2000). Remembering what happened: Memory errors and survey reports. In A.A. Stone et al. (Eds.), The Science of Self-Report: Implications for Research and Practice (pp. 29–47). Mahwah, NJ: Lawrence Erlbaum.

U.S. Conference of Mayors. (2001). A Status Report on Homelessness and Hunger in America's Cities. Washington, D.C.: U.S. Conference of Mayors.

Weinhardt, L.S., Carey, M.P., Maisto, S.A., Carey, K.B., Cohen, M.M., & Wickramasinghe, S.M. (1998). Reliability of the Timeline Follow-Back sexual behavior interview. Annals of Behavioral Medicine, 20(1), 25–30.

Wilcoxon, F. (1945). Individual comparisons by ranking methods. Biometrics, 1, 80–83.