THE DEVELOPMENT OF A PROBLEM BEHAVIOR ASSESSMENT TOOL

by

Kathy Montgomery, M.A.

A Dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

in the Department

of

Psychology

We accept this dissertation as conforming
to the required standard

[redacted]

Louis Costa, Ph.D.

[redacted]

Otfried Spreen, PH.D.

[redacted]

Michael Hunter, Ph.D.

[redacted]

Beverly Timmons, Ed.D.

[redacted]

Robert Willihnganz, Ph.D.

[redacted]

Martha Donnelly, M.D.

[redacted]

University of Victoria

ABSTRACT

Supervisor: Louis D. Costa, Ph.D.

The involvement of psychologists in the assessment and diagnosis of patients with possibly neurological disorders has a considerable history. More recently, psychologists have begun the provision of behavioral care and therapy to affected patients. The development of cognitive rehabilitation programs and centers intended to maximize the life skills of young head-injured patients is perhaps the most publicized of these efforts. Yet smaller numbers of psychologists are also trying to assist Long Term Care facilities in improving the behavioral care available to facility residents.

Psychological tests play a central role in assessment. Yet the test instruments available to date offer little benefit in the development of behavior management plans. This dissertation is about the development of an instrument that attempts to gather systematic information about the behaviors of Long Term Care (LTC) residents in a fashion that allows an initial clinical picture of behavior problems to be formed.

The instrument, named the Problem Behavior Assessment Tool (PBAT), was developed after frequent consultations with nurses regarding the clinical characteristics of problem behaviors observed in LTC residents. It contains two parts. The first part asks questions answered "yes" or "no"; the second part requires a further elaboration of existing problem behaviors, and delineates their "who, what, when and where" aspects. The PBAT contains three unique features not seen in other behavioral assessment tools: (a) it rates behaviors considered

to be problems by care staff, not by outside professionals, (b) it measures frequency and severity of problem behaviors via qualitatively different (multidimensional) items, not via a quantity of uniform items which yield a summed score; and, (c) it produces a preliminary picture of problem behaviors displayed by a resident.

The inter-rater and test-retest reliability of the PBAT are examined.

Examiners:

[redacted]

Louis Costa, Ph.D.

[redacted]

Otfried Spreen, PH.D.

[redacted]

Michael Hunter, Ph.D.

[redacted]

Beverly Timmons, Ed.D.

[redacted]

Robert Willihnganz, Ph.D.

Martha Donnelly, M.D.

[redacted]

TABLE OF CONTENTS

LIST OF TABLES

ACKNOWLEDGEMENTS

There are so many people who in one way or another made this dissertation possible. Just a few of them are included below. My thanks:

To my son Rod who taught me more about behavior (and misbehavior!) than anyone else, or any other way could have.

To my baby daughter Emily whose little hands were never in doubt about which computer key to press.

To my babysitter Lammy Andriessen whose loving care of my children meant I never needed to worry about their safety and happiness.

Especially to Roger whose support and kindness made this dissertation possible; who tamed the wild computers, and made the electricity behave.

To the senior management of Juan de Fuca Hospitals for their encouragement, and to their nurses for their patience and diligence.

To the Ruth Gardner Award, which presented the opportunity to do this project.

To my advisory committee for allowing me the time to work this out.

DEDICATION

To Roger

INTRODUCTION

In British Columbia, people requiring supervised living arrangements may be assessed as appropriate for Personal Care, Intermediate Care or Long Term Care. The amount of professional and non-professional supervision desirable increases in the order given. People assessed as requiring Personal Care need minimal assistance with activities of daily living and safety monitoring, but are otherwise independent. Those suitable for the three levels of Intermediate Care range from needing slightly more assistance than possible in Personal Care to those who are physically dependent and sometimes behaviorally disturbed. Most are elderly. Long Term Care (LTC) provides continuous professional supervision for people requiring special nursing procedures on a routine basis while not being acutely ill. Many LTC residents have considerable cognitive deficits.

Although facilities providing supervised living arrangements may provide adequate health monitoring and care, they are rarely anyone's first choice of abode. People entering such facilities may feel angry at their families for placing them, may be understandably depressed by their life situation, and may have a lifetime's accumulation of unresolved conflicts, personal difficulties and behavioral problems. Demented residents may inadequately perceive and comprehend their physical and social environment and respond inappropriately. Dementing illnesses may be accompanied by increased irritability, agitation and behavioral disturbance. Staff supervising LTC residents are faced

daily with the need to manage the difficult behaviors of a diverse resident group.

Behavioral disturbances range from mild problems that may do nothing more than irritate others, to severe disturbances, such as physical aggression, which jeopardize the safety of everyone present.

In the past, severe behavior disorders were suppressed by physical or chemical restraints, but it is now recognized that these methods are not solutions and that their use in elderly groups presents greater medical risks than in other adults. More recent management alternatives have been influenced by concepts developed for behavioral therapies. Group programs, such as Reality Orientation (Folsom, 1968) and Validation Therapy (Feil, 1967,1982), do not, however, focus on the specific behavior problems observed in individual residents, and their effect may not generalize to such problems. Psychologists might like to see programs tailored to fit individual LTC residents, but it is unrealistic to believe that there are resources available in the current health care system that could finance such care.

Since group programs are unlikely to change specific problem behaviors and individual programs are too costly, another option must be sought. The author proposes a two step model be used. First, identify specific problem behaviors and determine their relative frequencies. Although LTC staff members commonly refer to problem behaviors, their terminology is not precisely defined. Nor are frequency data routinely collected or available. Such information is essential. Second, develop a repertoire of management techniques specifically intended for use with the more frequently occurring

problem behaviors. Professional assessment can then lead to recommendations about which technique to implement, and the staff could be taught how to implement the technique.

Additionally, facility managements track the characteristics of their populations for future planning purposes. Staffing levels are negotiated with reference to nursing workload, and the specification of problem behavior management would further clarify the care burden. Were resulting programs to be tried and found effective, the benefit to nursing staff, residents and their visitors would be clear. The British Columbia Ministry of Health has recognized the value of such data in evaluating the need for psychogeriatric services (Cummings, Cummings, Titus, Schmelzle, & MacDonald, 1982).

An instrument allowing uniform collection of data specifically on problem behaviors is necessary to fulfill this first step. A survey of available instruments has failed to reveal a suitable one. To illustrate this point, some of the better known instruments are reviewed below.

PHYSICAL AND MENTAL IMPAIRMENT OF FUNCTION EVALUATION (PAMIE)

This scale, developed by Gurel, Linn, and Linn (1972), has 72 yes/no items divided into dimensions rating physical, psychological and social/personal functioning. It was developed from previous scales and was intended to allow raters (nurses) to make judgements of observable behaviors rather than requiring them to make inferences regarding those behaviors. Scale items include the following examples: "Most people would think him a mental patient," "Is distrustful and suspicious" and "Eats a regular diet". Pablo (1976) reported interrater reliability

figures ranging from .29 to .68 for varying parts of the scale, and internal consistency figures ranging from .70 to .91. The scale is generally considered adequate for the classification and description of severely impaired demented patients (Honingfeld, 1983).

NURSES OBSERVATION SCALE FOR INPATIENT EVALUATION (NOSIE)

This instrument is perhaps the most widely used behavioral rating scale intended to be completed by nursing personnel. It is comprised of either 30 or 80 items (depending on the version selected) used to rate the frequency of a specific behavior according to a 5 point scale (Honigfeld & Klett, 1965). Items were selected after investigation of a pool of 100 items gleaned from previous scales. Examples of scale items include: "Is messy in his eating habits", "Gets angry or annoyed easily," and "Is slow moving and sluggish". Though originally intended for use with an older schizophrenic population, it has been considered appropriate for geriatric samples. A unique feature is its inclusion of items regarding patient assets, not solely deficits. Its primary use is the assessment of treatment effects in psychopharmacological studies with psychotic or affectively disturbed patients. Seven dimensions identified through factor analysis are: social competence, social interest, cooperation, personal neatness, irritability, manifest psychosis, and psychotic depression. Interrater reliability among the factors ranges from .73 to .89 according to Honigfeld and Klett (1965).

SANDOZ CLINICAL ASSESSMENT - GERIATRIC (SCAG)

This instrument is one of the few multiple psychopathology rating scales available for specific use with elderly samples (Shader, Harmatz, & Salzman, 1974; Yesavage, Adey, & Werner, 1981). The 18

items were selected to best assess symptoms the test constructors thought pertinent to the status of mentally deteriorated elderly. The test covers four broad areas of impairment: cognitive impairment, mood and behavior, ability to cope with ordinary activities of everyday living, and somatic complaints. Ratings are made on a seven point scale. An item example is: "Hostility: verbal aggressiveness, animosity, contempt, quarrelsome, assaultive. Rate on impression at interview and patient's observed attitude and behavior towards others." This scale has been widely used in geriatric psychopharmacologic research and is considered valid and reliable (Venn, 1983). An inter-rater intraclass correlation of .75 is frequently cited for this tool (Shader, Harmatz, & Salzman, 1974), but it should be noted that this value was obtained by having the raters (psychologists or psychiatrists) watch the same clinical interview and not by conducting independent interviews as is the situation in actual clinical practice.

GERIATRIC RATING SCALE

A revision of the Stockton Geriatric Rating Scale (Meer & Baker, 1966), the Geriatric Rating Scale (Plutchik & Conte, 1970) has 28 items rated 0-2 and 3 unscored items. Its inter-rater (ward personnel) reliability is .87 to .94, and test-retest reliability based on a one year interval is in the mid .60's (Plutchik, Conte, Lieberman, Bakur, Grossman, & Lehrman, 1970). An item example is: "The patient knows the names of: 0= more than one member of the staff, 1= only one member of the staff, 2= none of the staff." Factor analysis reveals 3 dimensions: withdrawal/apathy, antisocial, disruptive behavior, and

deficits in activities of daily living (Smith, Bright, & McCloskey, 1977).

The above instruments are reasonably representative of those available. Existing tools differ in three major dimensions:

1. target group: community, demented, psychotic, affectively disturbed, or mentally retarded population,
2. content area: psychopathology, geriatric functioning, ward behavior, mental status, dementia and depression ratings, multifunctional assessment, etc., and,
3. purpose: choices are limited to classification and description or assessment of treatment effects.

No single existing instrument is suitable for the assessment portion of the management plan recommended in this proposal. Although instruments for the appropriate target group are available, they were devised for different purposes. Neither diagnostic classification, as it is usually formulated, nor assessment of treatment effects is the goal in the present study. Existing instruments lack a body of items (content) relating to the kinds of problem behaviors encountered in actual practice. In addition, it has become clear to the author that the important point in problem behavior assessment is not solely the determination of the presence of problem behavior, but rather, to determine the salience of specific individual behaviors according to their severity, frequency and manageability.

The goal of this dissertation is to investigate and develop a process for defining and quantifying particular problem behaviors displayed by LTC residents as perceived by nursing staff, in a manner

that reveals clinical characteristics. Unlike existing instruments, the purpose of this tool is to gather information that can assist in an initial clinical definition of problem behaviors, and document the "who, what, when, where and why" aspects of the problem behavior so that management strategies may be attempted.

RELATION TO NEUROPSYCHOLOGY

One could legitimately ask what relationship this dissertation has to neuropsychology. An explanation follows.

Neuropsychology began as a discipline concerned with identifying whether the clinical presentation of a patient was the result of an organic brain lesion. As its expertise developed, the focus shifted to determining the localization of such lesions, and the description of the deficit. More recently, neuropsychology has become increasingly involved in devising therapies intended to improve the capabilities of affected patients. Much of this effort is directed at producing programs for cognitive rehabilitation, especially for head injured patients. Fortunately, some programs are now seeking to improve general life skills and reduce problem behaviors, as demonstrated by presentations at the Neuropsychological Rehabilitation conference at the University of Victoria in 1986. Yet these patients need not be the sole recipients of neuropsychology's behavioral therapy endeavors. Consideration of projected population changes makes it clear that the proportion of people in older age groups is the most rapidly increasing segment, and that the incidence of brain diseases in older age groups is substantial. It would therefore seem prudent for neuropsychology to involve itself in examining therapy options appropriate for managing

problem behaviors exhibited by patients afflicted with brain diseases commonly seen in old age.

One might immediately counter the above line of thought by drawing attention to the differences in clinical characteristics between brain injured and older age patients. The former group is relatively young and their lesions are usually unchanging, while the later group possesses the combined deficits of reduced further life expectancy, and progressive disease. Two factors render such argument invalid. First, medical expenditures have increased dramatically, and may continue doing so in the future. Therapies which allow efficient and humane management of increasing numbers of dementing older adults will clearly be at a premium. Second, behaviorally impaired patients require care and supervision regardless of the clinical characteristics of their disease. Management of problem behaviors is an integral part of these patients' care, so why not develop expertise?

Elderly people with diagnosed brain diseases may remain living in the community, but those with progressive disorders are likely to eventually reside in institutions (Dellasega, 1986). Such facilities are perhaps the proper site to begin examining behavioral management strategies.

METHODOLOGY

Data collection for this dissertation was conducted at the Juan de Fuca Hospitals in Victoria, British Columbia. The hospital's four geographic locations contain seven nursing units, and have a total resident population of approximately 490 residents. Each nursing unit cares for approximately 70 residents who are accepted for admission only if they meet the eligibility requirements for Extended Care, (see Appendix One). Units are administered by a Director of the Nursing Unit who reports to the hospital senior management. Daily care is provided by trained nurse aides under the direction and supervision of registered nurses. Medical treatment is supplied either by the unit physician or by the residents' general practitioner.

The development of a tool that accurately reflects the view of problem behaviors seen by Long Term Care nurses involved several separate tasks. These included:

1. A formulation of a working model of problem behavior from nurses's point of view. This task provides an indication of the content domain a valid instrument must cover.
2. Another task was the examination of existing instruments.
3. A third task was the comparison of existing instruments with the view of nurses.
4. A brief review of psychometric properties of behavior rating scales was conducted.
5. The main tasks were the development of the current tool, and,
6. the investigation of its reliability.

The following section provides an outline of the development of the ten scales of problem behavior, which are collectively named as the Problem Behavior Assessment Tool (PBAT).

WHICH BEHAVIORS ARE PROBLEM BEHAVIORS?

The first task was to determine which behaviors were viewed as problems by Juan de Fuca nursing staff, and why they were problems. This task ensures a working understanding of the content domain instrument ought to cover. Staff meetings (inservice training programs) were conducted by the author on each unit for the day shift and in some cases for the evening shift as well. Questionnaires requiring staff to name three residents whose behaviors made providing care difficult and three residents whose behavior made care easy were completed. The nursing staff also specified which behaviors led residents to be considered easy or difficult. These judgements were to be made on the basis of behavior only, not according to physical disabilities, medical condition, etc., and some time was spend trying to ensure that nursing staff understood this instruction. Hospital charts of named residents were skimmed for behavioral descriptions, and Care Coordinators were interviewed. This process allowed developing an initial working model of both the behavioral tolerance of nursing staff, and the kinds of behavior a tool would have to assess.

WHAT RELEVANT INSTRUMENTS EXIST?

The second task was a critical examination of existing instruments. These instruments included, among others: Physical and Mental Impairment of Function Evaluation Scale (Gurel, Linn & Linn, 1972), Adult Personality Rating Scale (Kleban, Brody, & Lawton, 1971),

Sandoz Clinical Assessment - Geriatric (Shader, Harmatz & Salman, 1974, and Venn, 1983), Nurses' Observation Scale for Inpatient Evaluation (Honigfeld & Klett, 1965), and the Geriatric Rating Scale (Plutchik, Conte, Lieberman, Bakur, Grossman & Lehrman, 1970). Many of the instruments available for possible use with geriatric populations are intended to supply professionals with systematic information elicited from nursing personnel. These instruments generally incorporate items regarding sensory and physical abilities, and competence in the activities of daily living, among items concerning the more central focus of assessing behavioral dysfunction. Existing instruments appeared to be unsuitable for the present purposes due to the reasons outlined in the Introduction.

DO EXISTING TOOLS ASSESS THE BEHAVIORS CONSIDERED TO BE PROBLEMS?

The third task of comparing the nurse's view with that of existing instruments, was accomplished as follows. While it had become clear that an existing instrument would not adequately serve the purpose of specifically measuring problem behaviors independent of physical ability and competence at activities of daily living, perhaps the items regarding problem behaviors were suitable. To investigate this possibility, a trial instrument was constructed. Items were gleaned from various rating scales (those listed above), modified, and combined in a potentially usable form. Items were selected that seemed to correspond to problem behaviors included on the inservice questionnaires and mentioned in discussions with Care Coordinators; i. e., items that might address the content requirements of a valid

instrument. The selection was presented at a Nursing Practice Committee meeting for their perusal and feedback.

Committee members (frontline nurses) made the deficits they saw in the instruments quite clear. Some items did not directly relate to problem behavior as specifically exhibited by LTC residents and some items were written in a cumbersome format that clearly could not be rated quickly. Other items, or the response choices by which the rating scale items were answered, were not specific enough. The committee members were unanimous in their view that the items gleaned from existing instruments were irrelevant, difficult to rate and uninformative. The value of presenting this selection was the resulting refinement of the concept of the required instrument.

The major problem with existing instruments seemed to be that they aimed to quantify an umbrella construct of "pathology" defined as dysfunction in any of a number of abstract psychological variables, each measured by a body of fairly uniform and redundant items. Yet interviews with Care Coordinators indicated that their view of "pathology" was not compatible with the above model. Care Coordinators seem to place little or no emphasis on an umbrella construct, but rather recognized a number of distinct and separate problem behaviors, each being quite concrete in nature, and differing in salience according to the characteristics of an individual resident's expression. Existing instruments present a model of one superordinate category unifying multiple subcategories, each assessed by a body of unidimensional items. Care Coordinators appear instead to have a working model of multiple separate categories differing in salience

depending upon multidimensional features. In short, the model underlying existing instruments did not map onto the model Care Coordinators appear to have integrated.

TOOL DEVELOPMENT

The main task, development of the scales of the Problem Behavior Assessment Tool, began. The trial instrument described above was entirely abandoned. Inservice questionnaires were again examined and provided the basis for dimensions included in the PBAT. Discussions were held with Care Coordinators and Directors of the Nursing Units (DNU's) regarding the characteristics of these problem behavior dimensions and clinically relevant factors in management. It was clear that the importance of problem behavior is not just its presence or absence (which is what most published scales measure) but rather the frequency, severity and manageability of behaviors that do occur. A good instrument would have to assess both quantity and quality of problem behaviors. This crucial feature is lacking in existing scales.

Before proceeding with the development of a new psychological instrument, it is wise to increase ones familiarity with what is known regarding that type of instrument and conditions which foster the attainment of its best potential. For this reason, a review of behavior rating scales is included here.

BEHAVIOR RATING SCALES

The development and application of behavior rating scales is a well investigated area. A review of established principles which should be respected in the development of new behavior rating scales is

in order. These principles involve three main areas: choice of observers or raters, type of situation being rated, and properties of scales themselves. The areas mentioned here will be discussed below.

First, the literature on behavior rating scales clearly indicates that the choice of observers or raters has a definite impact on the quality of obtained ratings. This effect occurs primarily in two ways. The first has to do with the raters themselves; they bring with them to the rating situation their own personal orientations. Cronbach (1975) notes that raters are interpreters of what they see, and tend to overlook incidents which do not fit their interpretation. While all raters tend to be lenient in their ratings, the degree to which this occurs varies. The second way has to do with the rater's experience of the ratee (Bernardin & Beatty, 1984). Raters who have had more opportunity to observe the ratee will yield ratings of higher quality than will those whose exposure has been more limited. When the rated behavior corresponds to the rater's usual experience with the ratee, higher quality ratings are obtained. Raters are believed to be most discriminating when rating those individuals who are either high or low in the characteristic being examined.

Rating schemes are the conditions under which a rating is made. Some schemes involve observations of ratee's performance in a standardized setting whereas others are made according to usual daily performance. Preferences among rating schemes depend on the purpose of the rating, availability and quality of raters, and the potential for raters to gain by distorting ratings. Observations from daily life

tend to be superior to other methods, and are generally more economical.

One form of daily life observation, Critical Incident Methodology, developed by Flanagan (1949), involves recording incidents as they occur and using them to characterize the overall performance of ratees. In this method, raters have the responsibility to select incidents worth reporting, to report occurrences exactly as they happened, including significant preceding events and environmental conditions, and must endeavor to keep fact and interpretation as separate as possible. But this method may be of limited value in a systematic review of behavior as it requires extensive judgement and training on the part of the rater.

Behavior rating scales offer a more practical method of capturing the value of daily life observations. Smith and Kendall (1963) discuss the advantages of behavior rating scales at length. Their suggestions for ensuring that a behavior rating scale can attain its highest potential include: (a) guarantee that scale items refer to actual observed behaviors, (b) when developing scales, utilize the feedback of individuals who are as much like the intended raters as is feasible, (c) write items in the terminology of the raters and their environment, and (d) ensure consensus on the qualities and levels of behaviors being rated.

Properties of rating scales effect the level of effectiveness they are able to attain. Each item contained in a behavior rating scale should refer to a single variable and not incorporate multiple aspects of a variable in one rating since the rater may want to respond

differently to each aspect and then be in a quandary about choosing his response. The body of responses offered by the item for the raters selection should cover the range of likely and logical possibilities. Elaboration of these aspects of behavior rating scales are given in Allen and Yen (1988). Items may be formulated as either graphic ones such lines, or as ones which present the choice of multiple categories where response is marked with a check. Graphic scales effectively convey the idea of a continuum along which a rating may be made, and those with labels, either verbal or numerical, are preferable to those without labels because they reduce error caused by the rater forgetting the correspondence between rating point and its meaning.

Nunnally (1978) claims that the physical appearance of scales is actually among one of their least important aspects. The physical appearance of graphic items may show either discrete steps or may present continuous steps. Discrete steps are chosen if preventing responses from falling between steps is desirable; continuous steps are preferable when conveying a rating continuum is particularly important. Both vertical and horizontal scales have been used and they are considered equally effective. Vertical scales are reminiscent of other measurements made in a vertical fashion, such as thermometers, while horizontal scales mimic reading.

According to Nunnally (1978), the number of steps on a scale effects its reliability. As steps increase from two to twenty, reliability increases markedly at first until about seven steps, and little is gained thereafter. A large number of steps runs the danger of irritating raters to the point of carelessness. Scales may present

odd numbers of steps or even numbers of steps. Odd numbers of steps allow a neutral point, and the expression of individual response styles. The body of items on a scale should be balanced for "yea saying" and "nay saying" tendencies.

APPLICATION OF PRINCIPLES REGARDING BEHAVIOR RATING SCALES

The above brief review of behavior rating scale literature identified a number of principles relevant to the development of the current tool. A mention of their application will be made here. First, principles regarding rater selection indicate that the floor nurse who spends much of her work time in the same environment as a Long Term Care resident will be a better judge of his behavior than a physician who has expertise with a particular rating scale. Second, ratings made according to daily behavior are more advantageous and will be used in the PBAT. Critical Incident methodology offers assistance in the developmental process of devising a tool for present purposes. Third, properties of rating scales which enhance their effectiveness will be borne in mind when writing items for the current tool.

AN EARLY VERSION OF THE PBAT

A first version of the PBAT was written by incorporating what had been learned about the more frequent problem behaviors up to that time with knowledge of what appropriate behavior rating scales should include. This version was shown to a number of Care Coordinators (6) and DNU's (7). At the start of such sessions, the Care Coordinators and DNU's were asked to state five problem behaviors. Issues regarding content and face validity were covered. Their feedback allowed making the scale more closely approximate the manner in which Juan de Fuca

staff members actually talk about problem behavior and about basic nursing activities, and a comparison of their stated problem behaviors with those included on the PBAT helped to insure content validity. Two successive modifications followed, each viewed by the nursing staff and modified according to their feedback. This reiterative process by which the PBAT was developed resembles the Echo Technique of Bavelas (1967).

The validity of any new instrument should be considered. When the purpose of an instrument is to directly measure something, content validity is most relevant. According to Nunnally (1967), content validity is ensured by: (a) a representative collection of items, and (b) a sensible method of test construction. The PBAT was developed through repeated discussions with nursing care staff ranging from nurse aides to DNU's, and developed from their reports of behaviors that they labelled problem behaviors. The initial inservice training program meetings conducted on each unit had gathered a large body of specific examples of problem behavior, and names of specific residents who exhibited problem behavior. This body of information was the foundation on which the PBAT was developed, and allowed the selection of a representative set of items. Test construction has been described. Because of this construction process, the PBAT is considered to have a reasonable degree of content validity.

Tool development had brought awareness that geographic, architectural and staff differences among units were potentially great enough to render the PBAT invalid in the comparison of units. For example, geographic differences among units affect the threat a

wandering resident faces if he or she escapes from the building. One unit has a front exit door very close to a busy road intersection, whereas another unit is well removed from the road. Nurses in the first unit worry more if their resident gets out the exit door than do nurses in the second unit. Architectural differences include variables such as the existence of more than one dining room (which allows separation of residents with mealtime problem behaviors), and the presence of stairways and elevators. Staff differences are harder to define, but are mentioned due to apparent differences in attitude, approach and experience with behavioral problems.

The differences mentioned above made it necessary to ascertain whether DNU's viewed problem behaviors in a manner similar enough to allow using a single assessment system in all units. The DNU of each of the seven nursing units was asked to participate in a rating of the problem behaviors included on the preliminary PBAT. The DNUs received copies of the instrument prior to the session in order to allow them to be sufficiently acquainted with the tool. Sessions were conducted individually. DNU's were again asked to state five problem behaviors as a further check on content validity.

DNU's were requested to assign a score of 0-10 to each problem behavior on the PBAT according to how much of a problem it caused on their unit. Analysis by Concordance W shows that they rank order problem behaviors in a similar fashion and that the association of rankings is significant (W greater than .5 and p beyond .01, Table 1). This result suggests that despite observed geographic, architectural and staff differences, the DNUs nonetheless view problem behaviors in a

manner similar enough to warrant some confidence in comparing problem behaviors among units.

TABLE 1 RANK ORDERING OF PROBLEM BEHAVIORS

	CONCORDANCE W	P
Rummaging	.85	.00
Removing	.55	.00
Verbal abuse	.63	.00
Physical abuse		
to staff	.69	.00
to residents	.70	.00
Sexual misbehaviors	.63	.00
Meals	.53	.00
Wandering	.62	.00
Climbing out of bed	.73	.00
Noise	.77	.00
Miscellaneous items	.25	.04

PBAT DESCRIPTION

The current version of the PBAT (shown in its entirety in the next section) as it has evolved from the process described above is composed of two sections. The first section (Quick Form) has two parts containing questions asking whether or not a resident shows a particular kind of problem behavior. For Part 1, if the answer is no, nothing further is done. For Part 2 only, if the answer is yes, the rater is instructed to complete a corresponding scale from section two.

The second section contains ten separate Special Scales, which follow a uniform format. The first question requires a rating of problem severity as measured by bisecting a 15 cm. line, along which anchor points have been placed. The subsequent questions seek to delineate qualitative and quantitative components of the behavior as displayed by the individual resident, and are essentially "who, what, when, where" questions. Each section two scale is completed only when

the resident is identified in section one as showing the particular problem behavior.

Hence the PBAT is an instrument whose length of administration is determined by the number of problem behaviors observed in a given resident. Residents without problem behaviors are rated only on the Quick Form. Residents with one problem behavior are rated on the Quick Form plus the corresponding Special Scale from section two, and so on. The value of this format is the speed with which all residents on a unit can be rated, and the appeal of a brief tool to Care Coordinators, while still gathering data useful for both facility administration purposes and for planning resident management strategies.

The ten scales of problem behavior measure the frequency and severity of specific and concrete problem behaviors recognized by nursing staff as complicating the provision of quality care to LTC residents. It should be emphasized that the scales contain two unique features not seen in other behavioral assessment tools. The PBAT is unique in that: (a) it rates problem behaviors identified by staff, not selected by outside professionals; and, (b) it measures frequency and severity via qualitatively different (multidimensional) items, not via a quantity of uniform items which produce a summed score.

Following the development of the PBAT, an investigation of its inter-rater and test-retest reliability was necessary.

DESCRIPTION OF RELIABILITY PROCEDURES

The hospital administration was sufficiently pleased with the PBAT that they paid for replacement nurses in order to allow time for regular Care Coordinators to complete the PBAT for all residents without interruptions.

Site Recall that this investigation was conducted at the Juan de Fuca Hospitals in Victoria, British Columbia. The hospital's four geographic locations contain seven nursing units, most of which house approximately 70 residents, resulting in a usual total of approximately 490 residents.

Subjects Subjects were current residents of the Juan de Fuca Hospitals (N=463). The age range of residents is 32-110 years, with 45% over age 90 and 6% under age 70 (Mantle, 1985). Approximately 85% are female and 15% are male. Fifty percent are judged by the nursing staff as being continuously confused, 25% are periodically confused and 25% are rational all of the time. Residents meet criteria established for eligibility for Extended Care placement (see Appendix One).

Raters Raters were the two day shift Care Coordinators judged by the Director of the Nursing Unit as most familiar with the residents. A total of 15 pairs of raters from a pool of 28 nurses were involved. Raters had been working for the hospital an average of 8.89 years, and most had prior geriatric nursing experience. PBAT forms were completed by Care Coordinators for all residents, but some data were unusable due to admission less than two weeks previously (four residents), or death (one resident) between the two rater's sessions.

Data Collection Rating sessions were conducted with individual Care Coordinators. Following an introduction to the current PBAT, instructions on its proper completion, and time for answering any questions, PBAT's were completed for all residents in their care.

Approximately two weeks later, Care Coordinators were asked to re-rate four residents chosen in a pseudo-random fashion. This procedure resulted in the collection of data that allows inspection of inter-rater agreement and test-retest agreement.

DESCRIPTION OF STATISTICS USED

Data were analyzed with three types of statistics: 1. percent agreement, 2. Kappa, and 3. T equation. These statistics are described below, with the aid of the following model:

		Rater One		
		Y	N	
Rater Two	Y	A	B	A+B
	N	C	D	C+D
		A+C	B+D	A+B+C+D

The model is explained as follows. Raters can answer PBAT items either yes or no. Yes responses are coded "Y" above, and no responses are coded "N". Cell A represents "yes" responses by both rater one and rater two; cell D represents both raters responding "no". Cells B and C represent one rater responding "no" when the other rater has responded "yes".

1. Percent agreement Mean percent agreement is frequently used in the behavioral literature as a measure of reliability, and 80% has been suggested to be the minimum acceptable value (Hartmann, 1977). It is calculated as A + D/ A + B + C + D. There are a number of problems

with its use. The stability of the obtained percent agreement is greatly dependent on the sample size; the smaller the sample, the more easily is percent affected. Similar percent agreement figures do not necessarily indicate similar quality of observer agreement.

Methods for calculating percent agreement of behavior occurrence or non-occurrence have been proposed (Jensen, 1959) and continue to be advocated (Hartmann, 1977). These methods calculate percent only for cases where one or both raters claim a behavior to have occurred. Occurrence percent is equal to A / A + B + C, while non-occurrence percent equals D / D + B + C. These are thought to reflect accuracy of behavioral definition and thoroughness of rater training. Kratochwill and Wetzel (1977) particularly recommend the use of occurrence percent where the observed rate of behavior is very low. It should be noted, however, that these methods are affected by some of the same disadvantages as are the usual percent figures.

Percent agreement is affected by the observed frequency of rated behavior in that low frequencies predispose toward low percent agreement that the behavior occurred, and high frequencies toward high percent agreement that the behavior occurred (Yelton, Wildman, & Erickson, 1977). As a demonstration of this sensitivity, consider the situation where 10 observations are made. If the behavior of concern is of low frequency, it may occur twice, and two raters may indicate that it occurred twice, either agreeing or disagreeing on which two particular instances. Their agreement figures may range from 0 to 100 percent. If the behavior of concern is of higher frequency, one rater may indicate it occurred 6 times while the other rater indicates 8

times; they must agree on at least 4 instances, a minimum agreement of 40 percent.

Although percent agreement measures typically overestimate reliability, there are instances where there are no alternatives to their use. Variability is essential to the calculation of other reliability indices, and without it, percent agreement is the only choice. This is the situation when all cases are placed in one category. As Bartko and Carpenter (1976) state: "What is the reliability of an item which occurs very infrequently where (raters) agree on its absence? At present, a percent agreement descriptive statistic appears to be the only recourse" (pg. 309). Actually a statistic is meaningless here; a more reasonable approach is to simply describe what happened.

2. Kappa Kappa was proposed by Cohen in 1960 as an interrater reliability statistic for nominal scales that adjusts for chance agreements. It was designed to be a measure of agreement not association. It is defined as obtained agreement minus chance agreement, and, again referring to the above model proportions, it can be calculated as follows:

$$K = \frac{2(AD - BC)}{(A+B)(B+D)+(A+C)(C+D)}$$

Cohen's original formula was presented in a different fashion, and is shown here:

$$K = \frac{P(observed)-P(chance)}{1- P(chance)},$$

where P= proportions of agreement in the sample.

Kappa is appropriate for cases where scaling is nominal, i.e., placement in unordered categories. Its assumptions are: 1. units are independent, 2. categories are independent, mutually exclusive and exhaustive, and 3. judges operate independently.

Agreement is a special case of association. As mentioned above, ratings may be associated in that if one score is known, the other can be predicted. Agreement is the special case of association where both scores fall in identical categories. Hence, a set of ratings may show low agreement with high association, or may show high agreement and high association.

Kappa ranges from -1 for less than chance agreement, to 0 for exactly chance agreement, to +1 for perfect agreement.

Kappa is used primarily in two areas. One is the investigation of the reliability of psychiatric diagnoses (Spitznagel & Helzer, 1985), and the other is in the determination of interobserver reliability in behavioral research. For psychiatric studies, Grove et al. 1981 recommend .5 as a minimally acceptable kappa, whereas Hartmann (1977) states that .6 is the standard in behavioral research. For present purposes, .6 will be considered the minimally acceptable kappa.

Kappa is affected by base rate, and Grove et al. (1981) recommend that it not be reported when the base rate falls below 5%, as it is greatly lowered in instances of rare symptoms or infrequent occurrences. Kappa can not be calculated when both raters place all cases in the D cell above, i.e., agree completely that all cases belong

in the "no, no" cell. Kappa will be low both when 1. agreement is close to chance, and 2. when there is high agreement on low frequency of occurrence (Hartmann, 1977).

Kappa is restricted by the marginals. A maximum kappa can also be determined and the ratio of kappa to kappa maximum is again used to show how much of the marginally allowable agreement was obtained.

3. T Equation Tinsley and Weiss (1975) proposed a means of identifying whether agreement of ordinal scale data is of high, moderate or low degree. Their statistic, T, is calculated as follows:

$$T = \frac{N1 - NP}{N - NP}$$

N = number of subjects
N1= number of agreements
P = probability

T is similar to a phi correlation in that 0 equals chance agreement, 1 equals perfect agreement, positive values indicate above chance agreement, and negative values indicate less than chance agreement. T is appropriate for variables which are measurable on an ordinal scale.

Data analysis will present percent agreement and kappa since these statistics are appropriate for nominal scales. T values will be presented where measurement may be considered ordinal.

PROBLEM BEHAVIOUR ASSESSMENT TOOL

developed by:

KATHY MONTGOMERY, M.A.
PSYCHOLOGY PROJECT DIRECTOR

JUAN DE FUCA HOSPITALS
1450 Hillside Avenue
Victoria, B.C.
V8T 2B7

Tel: (604) 595 5722

JUAN DE FUCA HOSPITALS

PROBLEM BEHAVIOUR ASSESSMENT TOOL

INTRODUCTION

The PBAT is a rating scale designed to quantify problem behaviours occurring in Juan de Fuca residents in a manner similar to the way the Workload Assessment Tool quantifies nursing load.

The problem behaviours included on the rating scale were ones most frequently mentioned by Juan de Fuca Care Coordinators, Nurse Aides and Reactivation Aides on inservice questionnaires regarding resident behaviours that made the provision of care difficult.

The PBAT has two main components:

The first main component is entitled the Quick Form. The Quick Form contains two types of questions. The first type concerns irritating behaviours, and is simply answered "yes" or "no". The second type concerns more problematic behaviours and is also answered "yes" or "no", but the "yes" answer is accompanied by an instruction to complete a special scale (located in the second main component) regarding that problem behaviour.

The second main component is a collection of separate scales, each regarding a specific problem behaviour. These special scales correspond to problem behaviours from the Quick Form, and require specifying characteristics, such as severity, frequency, etc. Only scales for problem behaviours actually observed in a given resident are completed for that resident; the other scales are omitted.

.../...

-II-

By having these two main components, the PBAT is as short as possible while still being specific. For example, a resident who displays only the more minor irritating behaviours and no problem behaviours will require only the Quick Forms; whereas the resident who displays three problem behaviours will require the Quick Form plus the corresponding three special scales.

The Quick Form should be answered completely before moving on to the special scales.

PROBLEM BEHAVIOUR ASSESSMENT TOOL - QUICK FORM

Resident's Name: _____ Rater's Name _____

Unit: _____ Side: A B C

Date: _____

Questions in this section are somewhat general and broad, as for example, "Is the resident verbally abusive ?" These questions should be answered according to the characteristic, typical behaviour of the resident as you have come to know him or her. Answer according to typical behaviour, not according to one unusual incident. Consider only the resident's current behaviour in the last four weeks.

For example, if the question is "Is the resident verbally abusive"? and the resident you are rating becomes nasty, hostile and swears on a regular basis, you would answer yes. Conversely, if you are rating a resident who required an unpleasant treatment and responded by being verbally abusive for about the only time you are aware of, you would answer
no, it is not typical of the resident to be verbally abusive.

All questions on the Quick Form should be answered before filling out the special scales.

Component One - Part One
Mark an answer for each of the questions below.

---

1. Is the resident considered demanding, "spoiled" or manipulative by many staff members?

	Yes	No
	___	___

---

2. Does the resident insist on an excessive amount of assistance considering his or her mental and physical abilities?

	Yes	No
	___	___

---

3. Is the resident uncooperative or resistant to:

	Yes	No
(a) morning care, h.s. care	___	___
(b) toileting	___	___
(c) nutrition	___	___
(d) transferring	___	___
(e) medication administration	___	___
(f) treatments	___	___

---

4. Is the resident generally irritable, grouchy unpleasant?

Yes	No
___	___

5. Does the resident make complaints that the staff generally believe to be unjustified?

Yes	No
___	___

Part Two

Mark an answer for each of the questions below.
There is a row of numbers printed at the end of all of the questions. Circle the numbers corresponding to the scale numbers for questions answered yes.

Does the resident rummage thru other's things ? (rummage means going "going through" others things - looking through drawers, cupboards, etc).

Yes - (answer questions on scale 1)
No

Does the resident remove things inappropriately ? (remove means takings away from where they are kept when the resident does not own them or need to use them).

Yes - (answer questions on scale 2)
No

Is the resident verbally abusive ?

Yes - (answer questions on scale 3)
No

Is the resident physically abusive to staff ?

Yes - (answer questions on scale 4)
No

Does the resident physically abuse other residents ?

Yes - (answer questions on scale 5)
No

Is the resident involved in sexual misbehaviours ? (examples: getting fresh, open masturbation, public exposure, unacceptable sexual comments ?

Yes - (answer questions on scale 6)
No

(iii)

Are there problems concerning the resident's meal time behaviour ?

Yes - (answer questions on scale 7)
No

Does the resident try to climb out of bed even though he/she lacks the ability to do so safely ?

Yes - (answer questions on scale 8)
No
Not applicable

Does the resident resist bed restraints ?

Yes - (answer questions on scale 8)
No
Not applicable

Wandering is defined as going to, or being in, a place where the individual resident is not allowed. It implies getting into some kind of trouble in an out-of-bounds area.

Wandering does not include aimless moving about within proper areas.

Does the resident wander ?

Yes - (answer questions on scale 9)
No

Does the resident create disruptive unnecessary noise ?

Yes - (answer questions on scale 10)
No

Circle scales to be completed :

1 2 3 4 5 6 7 8 9 10

Now answer questions on the scales circled above.

INSTRUCTIONS FOR SPECIAL SCALES

Complete only the special scales which correspond to questions answered "yes" on the Quick Form.

The special scales contain questions written in two formats. One format presents a question which is answered by checking off one or more items listed below it. The second format presents a question followed by a line which has words printed along it. These questions are answered by putting a vertical slash thru the line. DO NOT CIRCLE THE WORD. Scores for these questions are determined by measuring the line from its far left to the slash mark.

The first question on each scale asks how serious a problem a given behaviour is. Answers range from "not a problem" to wishing the resident could be transferred". Answers in the "severe" or "wishing the resident could be transferred" areas reflect the strain placed on a unit as a whole by a resident's difficult behaviour. Answers in these areas do not label a resident as bad; they given an honest reflection of the difficulty encountered in trying to manage some problem behaviours displayed by residents despite everyone's best efforts.

Scale 1

RUMMAGING*

* rummage means - "going thru" other's things, e.g. looking thru drawers, without taking things away.

1. How serious a problem is the resident's rummaging ?

Not a problem	Mild	Moderate	Severe	Wish resident would be transferred because rummaging is so severe.

2. Where does the resident like to rummage? Mark any that apply

_____ In other resident's things
_____ In innocuous things belonging to the unit
_____ In potentially dangerous areas, e.g., med cart, cleaning supplies

3. Does the resident stop rummaging readily at staff's request or intervention?

Stops readily	May stop	Does not stop without staff's insistent effort

4. How frequently does the resident rummage?

Every day	Nearly every day	Several times a week	Several times a month

5. What do you think prompts the resident to rummage? Mark any that apply.

_____ Boredom
_____ Confusion
_____ Is trying to find something to do
_____ Don't know
_____ Is trying to find a particular thing; e.g., candy
_____ Is looking for something, though can't say what
_____ Other

Comments

Scale 2

REMOVING

* remove - means that the item is taken away from where it's kept.

1. How serious a problem is the resident's removing ?

Not a problem	Mild	Moderate	Severe	Wish resident would be transferred due to removing

2. How frequently does the resident remove things ?

Every day	Nearly every day	Several times a week	Several times a month

3. Does the resident remove things that :

____ are readily available (e.g., takes things off beside table or w/c tray).

____ require extra effort to obtain ? (e.g., takes things out of bedside tables, closets, etc).

4. What does the resident remove ? Mark any that apply.

____ Edible items, e.g., food, drink, candy

____ Other resident's personal possessions, e.g., clothing, cosmetics, grooming aids, teeth

____ Items belonging to the unit

____ Other (please specify if possible)

5. Does the resident have the same kind of object as he has removed? Mark any that apply.

____ Yes
____ Pretty much
____ No
____ Don't know

6. What does the resident do with things he or she has removed? Mark any that apply.

____ Keeps them openly, or leaves them just anywhere
____ Hides them
____ Stashes them in a particular place, e.g. stashed in large purse
____ Eats edible items
____ Uses them dangerously
____ Don't know

Comments

Scale 3

VERBAL ABUSE

1. How serious a problem is the resident's verbal abuse ?

Not a problem	Mild	Moderate	Severe	Wish resident would be transferred due to verbal abuse

2. What is the nature of the verbal abuse? Mark any that apply.

_____ Swears and curses, e.g. hell, damn, etc
_____ Obscene and vulgar language, e.g., words with a sexual meaning
_____ Makes verbal threats, accusations
_____ Name calling
_____ Harassment, belligerency, tongue lashing
_____ Other (please specify)

3. How frequently is the resident verbally abusive?

Every day	Nearly every day	Several times a week	Several times a month

4. Whom does the resident direct verbal abuse towards? Mark any that apply.

_____ Abuse is not directed towards anyone specifically - resident just lets out verbal abuse,could even be alone
_____ Directed toward staff
_____ Directed toward other residents
_____ Other (please specify)

5. Where is resident when being verbally abusive ? Mark any that apply.

_____ Private areas, e.g. bedroom, bathroom
_____ Common areas, e.g. dining room, hallways

6. When is the resident verbally abusive ? Mark any that apply.

_____ When receiving daily physical assistance, e.g. dressing, toileting
_____ When receiving a treatment
_____ When in pain or in physical discomfort
_____ During social interactions
_____ Other (please specify)

7. How loud is the verbal abuse?

_____ Audible in room only
_____ Audible in room & same hall only
_____ Audible throughout unit
_____ Audible outside building

Continue on to next page/.....

8. Is it generally possible to limit the volume of the resident's verbal abuse such that it is only audible in one room?

_____ The resident's volume can be limited, or is limited
_____ The resident's volume cannot be limited

Comments

Scale 4

PHYSICAL ABUSE TOWARD STAFF

1. How serious a problem is the resident's physical abuse of staff? Mark any that apply.

Not a problem	Mild	Moderate	Severe	Wish resident would be transferred due to physical abuse to staff

2. In what way(s) is the resident physically abusive to staff?

_____ Minor - e.g. spits, pushes, scratches, pulls hair
_____ Moderate - e.g. bites, kicks, slaps, twists arm
_____ Severe - e.g. tries to choke, punches, uses weapon

3. How frequently is the resident abusive?

Every day	Nearly every day	Several times a week	Several times a month

4. Has the resident caused WCB injury to a staff member within the last year?

_____ Yes
_____ Not to my knowledge

5. What is the resident's size ?

Height: _____

Short Average Tall

Weight: _____

Light Average Heavy

6. What are the resident's physical abilities ?

Strength: _____

Weak Average Strong

Mobility: ____ Walks without assistance
____ Uses walker without assistance
____ Wheelchair (moved without assistance)
____ Bedridden, unable to move about (by any means) without assistance.

Comments

Scale 5

PHYSICAL ABUSE DIRECTED TOWARDS OTHER RESIDENTS

1. How serious is the resident's physical abuse of other residents ?

Not a problem	Mild	Moderate	Severe	Wish resident would be transferred due to physical abuse to other residents

2. How frequently has the resident actually abused another resident?

Every day	Nearly every day	Several times a week	Several times a month

3. Who is usually responsible when this resident is involved in an altercation? Mark any that apply.

_____ The two residents share responsibility
_____ The resident being rated is responsible
_____ The resident threatened or abused is responsible
_____ Don't know or can't really say

4. How serious is the nature of the physical abuse? Mark any that apply.

_____ Minor
(e.g) small push to w/c, pinch, scratch
_____ Major
(e.g) punch, kick
_____ Potentially lethal
(e.g) chokes

5. What is the consequence of the abuse for the recipient?

_____ Emotionally upset, without physical consequence
_____ Bruise, scratch or similar result
_____ Medical care required
_____ Sent to acute hospital

Comments

Scale 6

SEXUAL MISBEHAVIOURS

1. How serious are the resident's sexual misbehaviour(s):

Not a problem	Mild	Moderate	Severe	Wish resident could be transferred due to sexual misbehaviours

HOW FREQUENTLY .....

2. Does the resident make unacceptable sexual comments:

To other residents:

Every day	Nearly every day	Several times a week	Several times a month	No

To staff members:

Every day	Nearly every day	Several times a week	Several times a month	No

3. Does the resident try to "get fresh" with staff members e.g. tries to touch breasts, pinch bottoms, kiss, etc. ?

Every day	Nearly every day	Several times a week	Several times a month	No

4. Does the resident try to touch other residents in unwelcomed and inappropriate "fresh" ways ?

Every day	Nearly every day	Several times a week	Several times a month	No

5. Does the resident expose his/her body in inappropriate places (e.g., hallway) ?

Every day	Nearly every day	Several times a week	Several times a month	No

6. Does the resident masturbate openly?

Every day	Nearly every day	Several times a week	Several times a month	No

Comments

Scale 7

MEAL BEHAVIORS

1. How severe are mealtime misbehaviours ?

Not a problem	Mild	Moderate	Severe	Wish resident could be transferred because of mealtime behaviour problems

2. Mark any that apply :

_____ The resident is noisy.

_____ The resident is sometimes abusive during meals.

_____ The resident's behavior is unpredictable.

_____ The resident does unacceptable things with food or utensils, e.g. throws, spits, plays, drops.

_____ These things are done in a way that might hurt someone, e.g. throws hot liquid.

_____ A staff member must devote their sole attention to feeding this resident because of his/her behavior.

_____ The resident must be cajoled into eating, (more so than the typical resident).

_____ The resident refuses to eat or drink enough.

_____ The resident eats too much.

_____ The resident eats, bites or chews things that he/she shouldn't.

_____ The resident is not fed in a DR because of his or her behavior.

_____ Other residents in the same DR don't want this resident at their table.

_____ The resident disturbs other residents eating in the same room.

Scale 8

CLIMBING OUT OF BED

1. Is it typical of this resident to try to climb out of bed ?

   ____ No, go on to #2 on this page, Bed Restraint
   ____ Yes, which of the below is the result when the resident tries to climb out of bed ?

      ____ will fall, is unable to stand or transfer;
      ____ may fall, is barely able to stand or transfer;
      ____ is able to get out of bed, but it would be better if he/she were assisted
      ____ fully able to get out of bed, assistance not required or desirable

2. Bed restraint: which one of the below is most correct ?
   ____ is not necessary
   ____ is not used because resident is so resistive, he or she might be injured
   ____ is used and mildly resisted; (removal not indicated)
   ____ is used and tolerated by resident

Scale 9

WANDERING

Wandering is defined as going to, or being in, a place where the individual resident is not allowed. It implies getting into some kind of trouble in an out-of-bounds area.

Wandering does not include aimless moving about within proper areas.

1. How serious a problem is the resident's wandering ?

Not a problem	Mild	Moderate	Severe	Wish resident could be transferred due to wandering

2. How frequently does the resident wander:

Into other resident's rooms ?

Every day	Nearly every day	Several times a week	Several times a month	Never or almost never

In other areas on the unit ?

Every day	Nearly every day	Several times a week	Several times a month	Never or almost never

Off the unit, but in the building ?

Every day	Nearly every day	Several times a week	Several times a month	Never or almost never	N/A

Outside the building ?

Every day	Nearly every day	Several times a week	Several times a month	Never or almost never	N/A

2. On which shifts does the resident wander?

Days	Evenings	Nights
____	____	____

3. How difficult is it to prevent or limit this resident's wandering? Choose one of the below

____ less difficult than most other wanderers
____ similar to most wanderers
____ more difficult than most wanderers

Continue on to next page/...

4. Once the resident has wandered, what is his/her response to staff attempts to return him/her to a proper area?

_____ cooperates
_____ makes verbal objections only
_____ physically resistive to return
_____ physically aggressive to staff

5. Is this resident wandering:

_____ in a wheelchair ?
_____ with a walker ?
_____ by walking ?

Comments

Scale 10

NOISE

(Verbal abuse has a separate page and is not included here).

1. How serious a problem is unnecessary noise created by this resident ?

Not a problem	Mild	Moderate	Severe	Wish resident could be transferred due to noise

2. How frequently does this resident create disruptive unnecessary noise?

Every day	Nearly every day	Several times a week	Several times a month

3. On what shifts does this occur?

Days	Evenings	Nights
____	____	____

4. What is the nature of the noise, and how loud is it?

	Audible		
	In resident's room only	In restricted part of unit only	Throughout unit
repetitive words (e.g. "help me! help me!")	____	____	____
yelling, screaming, moaning	____	____	____
crying	____	____	____
other noisy behavior	____	____	____

5. When the resident is being noisy, is the noise continuous or intermittent ? Choose one.

The noise is usually intermittent ____

The noise is usually fairly continuous ____

6. When the resident is being noisy, are you able to quiet them? Choose one.

Usually ____

Sometimes ____

Rarely ____

Comments

RESULTS AND CONCLUSIONS

The following sections outline the results of the evaluation of various parts of the PBAT. To begin with, the first part of the Quick Form is examined, and a table summarizing obtained results is provided, (Table 2). The results of the second part of the Quick Form are summarized in Table 3, but the narrative for this part of the Quick Form accompany the results given for each Special Scale. Then each of the Special Scales is investigated, looking at the questions and their items. A brief discussion of severity, frequency and other parts of each scale follows. A second version for a Problem Behavior Assessment Instrument is then given. Consideration of the diversity of rater pair performances is presented last.

Since the results are detailed and cumbersome, an outline showing subsequent sections is provided here in hopes that it will forewarn the reader and hence aid him in retaining an orientation to the material.

FIRST PART OF THE QUICK FORM
- INTER-RATER RELIABILITY
- TEST-RETEST RELIABILITY
- SUMMARY

SECOND PART OF THE QUICK FORM

SPECIAL SCALES
- RUMMAGING
- REMOVING
- VERBAL ABUSE
- PHYSICAL ABUSE TO STAFF
- PHYSICAL ABUSE TO OTHER RESIDENTS
- SEXUAL MISBEHAVIORS
- MEAL BEHAVIORS
- BED
- WANDERING
- NOISE

SEVERITY

FREQUENCY

OTHER LINE BISECTIONS

COMPARISON OF RATERS AND RATER PAIRS
INTER-RATER RELIABILITY OF RATER PAIRS
TEST-RETEST RELIABILITY OF RATERS THEMSELVES

Now, to the results.

FIRST PART OF THE QUICK FORM

Results are summarized in Table 2.

INTER-RATER RELIABILITY

A. Percent Agreement

Mean percent agreement indicates the total percent of times both raters agree on combined presence and absence of a problem behavior in the sample (A + D / A + B + C + D). For questions on the First Part of the Quick Form, the range is 79-89%. Nine questions in Part One exceed 80%.

Occurrence percent agreement indicates the percent of times the pairs agree on the presence of a problem behavior relative to the number of times either of the raters said it occurred (A / A + B +C). The range is 23-51%. It must be realized that occurrence percent agreement is apt to be low when the frequency of rated behavior in the sample is low (Hartmann, 1977) as is the case for the PBAT ratings.

Non-occurrence percent agreement indicates the rate at which it can be agreed that problem behavior is absent. The range is 66-84%, emphasizing the low frequency of specific problem behaviors.

B. Kappa

Kappa ratios were calculated. Two questions yielded kappa ratios over .6.

TEST-RETEST RELIABILITY

A. Percent Agreement

Mean percent test-retest agreement for questions on the first subjective, attitudinal section of the PBAT ranged from 80 to 90%. All are at 80% or above.

Occurrence percent ranged from 33-62%.

Non-occurrence percent ranged from 56 to 79%.

B. Kappa

Seven questions had a kappa ratios above .70.

SUMMARY

The questions included in the first subjective and attitudinal section of the PBAT had clinically relevant reliability, demonstrated by acceptable percent agreement figures. One question achieved a mean percent agreement of 79% for inter-rater reliability, but reached 80% for test-retest reliability.

Occurrence percent figures were higher for test-retest comparison than inter-rater comparison, indicating that raters are better able to agree with their own judgements than those of their partner.

Kappa ratio values, which examine positive pathology, were sufficiently higher in the test-retest situation. Instead of two acceptable values, seven were acceptable.

The questions of the First Part of the Quick Form were reliable enough for use in screening residents into positive and negative pathology groups.

Consideration of the differences between results from the two rating conditions allows one to question the impact of rater attitude and could suggest that rater attitude toward individual residents may differ markedly between members of the rating pair. If this is true, it would affect not only their rating performance, but quite possibly their care of the individual resident.

SECOND PART OF THE QUICK FORM

These results are summarized in Table 3. Since performance on the second part of the Quick Form precipitates completing the corresponding Special Scale, the narrative for these results is used to preface the results of each of the Special Scales individually.

SPECIAL SCALES

The next section examines the individual Special Scales. (Agreement statistics are summarized in Table 3 for convenience). Statistics supporting the reported results of each question on the Special Scales are given in Appendices Three and Four.

TABLE 2 AGREEMENT STATISTICS FOR THE FIRST PART OF THE QUICK FORM

QUESTION	YY	%	NY YN	%	NN	%	EP	DP	MEAN %	K	KMAX	KRATIO
INTER-RATER CONDITION												
1	58	13	56	12	349	75	51	37	88	.60	.89	.68
2	24	5	51	11	388	84	32	57	89	.42	.85	.50
3	66	14	93	20	304	66	42	38	80	.46	.89	.51
4	32	7	93	20	338	73	26	54	80	.29	.87	.33
5	55	12	74	16	334	72	43	41	84	.50	.82	.61
6	29	6	95	21	339	73	23	56	79	.26	.81	.32
7	36	8	77	17	350	76	32	52	83	.39	.85	.45
8	30	6	79	17	354	76	28	55	83	.33	.89	.37
9	36	8	88	19	339	73	29	52	81	.34	.97	.35
10	30	6	56	12	377	81	35	53	88	.45	.78	.57
TEST-RETEST CONDITION												
1	26	22	16	13	78	65	62	25	87	.67	.84	.80
2	9	7	18	15	93	77	33	52	85	.42	.74	.56
3	32	27	21	17	67	56	60	22	83	.62	.73	.85
4	28	23	20	17	72	60	58	25	83	.62	.85	.73
5	16	13	21	17	83	69	43	39	83	.49	.88	.56
6	16	13	24	20	80	67	40	40	80	.45	.77	.58
7	21	17	15	13	84	70	58	29	88	.66	.89	.74
8	16	13	18	15	86	72	47	38	85	.55	.75	.73
9	23	19	22	18	75	63	51	31	82	.56	.68	.82
10	13	11	12	10	95	79	52	38	90	.63	.94	.67

LEGEND
YY= number of times both raters answer "yes"
NYYN= number of times one rater answers "yes" and the other rater answers "no"
NN= number of times both raters answer "no"
EP= effective percent, occurrence percent
DP= mean percent minus effective percent
K= kappa KMAX= kappa max

TABLE 3 AGREEMENT STATISTICS SUMMARY FOR THE SECOND PART OF THE QUICK FORM

QUESTION	YY	%	NY YN	%	NN	%	EP	DP	MP	K	KMAX	KRATIO
SPECIAL SCALES - INTER-RATER CONDITION												
11	6	1	11	2	446	96	35	62	98	.51	.96	.53
12	8	2	14	3	441	95	36	61	97	.52	.79	.65
13	45	10	57	12	361	78	44	44	88	.54	.91	.59
14	54	12	62	13	348	75	47	40	87	.55	.96	.58
15	14	3	23	5	426	92	38	57	95	.52	.81	.64
16	5	1	9	2	449	97	36	62	98	.52	.73	.71
17	63	14	113	24	287	62	36	40	76	.37	.85	.43
18	18	4	30	6	415	90	38	56	94	.51	.97	.53
19	2	0	14	3	447	97	13	84	97	.21	.89	.23
20	10	2	27	6	426	92	27	67	94	.40	.71	.56
21	55	12	54	12	354	76	50	38	88	.60	.97	.62
SPECIAL SCALES - TEST-RETEST CONDITION												
11	18	15	5	4	97	81	78	18	96	.85	.91	.94
12	20	17	6	5	94	78	77	18	95	.84	.89	.94
13	45	38	16	13	59	49	74	13	87	.73	.87	.84
14	29	24	20	17	70	59	59	24	83	.62	.81	.77
15	23	19	10	8	87	72	70	22	92	.77	.86	.89
16	2	2	6	5	112	93	25	70	95	.38	.38	1.00
17	32	27	22	18	66	55	59	22	82	.60	1.00	.60
18	19	16	8	7	93	77	70	23	93	.78	.95	.83
19	2	2	10	8	108	90	17	75	92	.25	.70	.35
20	17	14	13	11	90	75	57	33	89	.66	.76	.86
21	46	38	11	9	63	52	81	10	91	.81	.92	.89

LEGEND
YY= number of times both raters answer "yes"
NYYN= number of times one rater answers "yes" and the other rater answers "no"
NN= number of times both raters answer "no"
EP= effective percent, occurrence percent
DP= mean percent minus effective percent
K= kappa KMAX= kappa max

RUMMAGING

INTER-RATER CONDITION

Twenty-three rummaging inventories were completed. Twelve were for six residents (1.3%) said by both raters to be rummagers, and 11 (2.4%) were for residents said to rummage by one rater. The remaining 446 residents (96.3%) were rated as not rummaging by both nurses. Percent agreement for the full sample of 463 residents was 97.6%. Kappa ratio was .53.

TEST-RETEST CONDITION

Twenty-two inventories were completed, including 18 (82%) agreements and 4 (18%) disagreements. Mean percent agreement based on the retested sample of 120 was 96%. Kappa ratio was .93.

1. HOW SERIOUS A PROBLEM IS THE RESIDENT'S RUMMAGING?

A. INTER-RATER RELIABILITY

Pairs showed little correspondence in their use of the severity scale, either as a five-point scale (T= .11, NS) or a two point scale, (T= .33, NS). The kappa ratio was .25.

B. TEST-RETEST RELIABILITY

Raters showed no reliability in their use of the severity scale, either as a five-point scale (T= -.07, NS) or as a two point scale, (T= .20. NS). In fact, rummaging severity has a mean of 62.1 in the inter-rater condition and a mean of 46.20 in the test-retest condition. The kappa ratio was below the criterion.

2. WHERE DOES THE RESIDENT LIKE TO RUMMAGE?

A. INTER-RATER RELIABILITY

Pairs showed very good agreement when rating whether residents rummage in innocuous unit possessions (83%, kappa ratio 1.00), but not when rating whether residents rummage in other resident's things or in potentially dangerous areas.

B. TEST-RETEST RELIABILITY

Agreement on the three choices to this question was 83%. Raters showed high agreement when rating whether residents rummage in innocuous unit possessions (kappa ratio 1.00) and in other resident's things (kappa ratio 1.00), but not when rating whether they rummage in potentially dangerous areas.

3. DOES THE RESIDENT STOP RUMMAGING READILY AT STAFF'S REQUEST OR INTERVENTION?

A. INTER-RATER RELIABILITY

Pairs did not agree about their ability to stop a resident who was rummaging, (17%, kappa ratio below .6). Either the scale cannot be used properly, or they actually differ in their ability to stop a rummager.

B. TEST-RETEST RELIABILITY

Agreement was 60%, while the kappa ratio was close to criterion at .57. If pair members differed in their ability to stop a rummager, one could expect to see a higher degree of agreement on the test-retest evaluation.

4. HOW FREQUENTLY DOES THE RESIDENT RUMMAGE?

A. INTER-RATER RELIABILITY

Pairs showed insufficient agreement about the frequency of rummaging, in that the percent value was 67% though the kappa ratio was 1.00.

B. TEST-RETEST RELIABILITY

Rater's agreement about the frequency of rummaging was again below 80%,but the kappa ratio was .79.

5. WHAT DO YOU THINK PROMPTS THE RESIDENT TO RUMMAGE?

A. INTER-RATER RELIABILITY

Pairs did not indicate agreement in their ideas of why a resident rummages. All but one agreement was below 80%, and no kappa ratio was above .6.

B. TEST-RETEST RELIABILITY

Raters agreed about most possibilities concerning what might prompt residents to rummage, with four agreements above 80% and kappa ratios equalling 1.00 in two instances and exceeding .6 in a third.

SUMMARY

Agreement about whether or not a resident rummages can be shown on the PBAT. Kappa ratios for the test-retest frequency of rummaging were acceptable for raters. Raters also maintain their opinions about what prompts rummaging. Other clinical characteristics of rummaging behavior can not be judged reliably.

The following table gives descriptive information regarding items from the Rummaging Scale. It is based only on those instances where a resident was rated as rummaging either by both raters (inter-rater) or on both occasions by the same rater (test-retest). Results of subsequent scales contain similar tables.

TABLE 4 DESCRIPTIVE INFORMATION - RUMMAGING

		MEAN	STD DEV	MINIMUM	MAXIMUM
INTER-RATER					
SEVERITY		60.92	24.46	27	95
STOP		83.41	48.11	13	136
FREQUENCY		69.00	45.25	8	150
TEST-RETEST					
SEVERITY	T1	62.10	24.97	27	95
	T2	46.20	21.89	9	73
STOP	T1	86.20	45.48	14	136
	T2	87.90	37.49	20	136
FREQUENCY	T1	63.40	47.88	8	150
	T2	52.20	52.14	10	144

REMOVING

INTER-RATER CONDITION

Thirty removing forms were completed. This group included 16 forms for 8 residents (1.1%) said by both raters to remove and 14 forms for residents (3%) said to do so by only one rater. Raters agreed that 441 residents (95.2%) did not remove things inappropriately. Mean percent agreement was 97%. Kappa ratio was .65.

TEST-RETEST CONDITION

Twenty-five removing forms were completed. Twenty (80%) were agreements with the rater's previous judgement and five (20%) were disagreements. Mean percent agreement for the retested sample was 95%, while the kappa ratio was .94.

1. HOW SERIOUS A PROBLEM IS THE RESIDENT'S REMOVING?

A. INTER-RATER RELIABILITY

Pairs did not agree about the severity of this problem either as a two-point ordinal scale, (T= .50, p NS), nor as a five-point ordinal scale, (T= -.17, p NS). Kappa ratio does not exceed .6.

B. TEST-RETEST RELIABILITY

Raters agreed about the severity of this problem as a two point scale, (T= .80, p <.001), but not as a five-point scale, (T= .27, p NS). Kappa ratio does not exceed .6.

2. HOW FREQUENTLY DOES THE RESIDENT REMOVE THINGS?

A. INTER-RATER RELIABILITY

Pairs did not agree about the frequency with which removing occurs, (T= .33, p NS). Kappa ratio was less than .6.

B. TEST-RETEST RELIABILITY

Raters did agree about the frequency with which removing occurs, in that the T value was significant at .001, but the kappa ratio was .54.

3. DOES THE RESIDENT REMOVE THINGS THAT ARE READILY AVAILABLE OR REQUIRE EFFORT?

A. INTER-RATER RELIABILITY

Pairs did not agree about the availability of things that are removed.

B. TEST-RETEST RELIABILITY

Raters did agree about the availability of things that were removed, (agreement was 80% and kappa ratios range from .60 to .74).

4. WHAT DOES THE RESIDENT REMOVE?

A. INTER-RATER RELIABILITY

Pairs agreed about three of the four categories of items removed by residents. Agreement ranged from 75 to 88%, and kappa ratios were 1.00.

B. TEST-RETEST RELIABILITY

Raters were able to agree quite well about the category of items

removed by residents. Percent agreement ranged from 75 to 100 percent and kappa ratios were above .7 for three of the four choices.

5. DOES THE RESIDENT HAVE THE SAME KIND OF OBJECT HAS HE HAS REMOVED?

A. INTER-RATER RELIABILITY

Pairs agreed that they did not know whether residents have an item like one they remove,i.e., they showed good inter-rater reliability when marking the answer "don't know" to this question.

B. TEST-RETEST RELIABILITY

Raters also agreed that they did not know whether residents have an item like one they remove,i.e., they showed complete reliability when marking the answer "don't know" to this question. This dimension is probably not particularly relevant.

6. WHAT DOES THE RESIDENT DO WITH THINGS HE HAS REMOVED?

A. INTER-RATER RELIABILITY

Pairs agreed with each other when answering 1. they do not know what happens to items, (88%) and 2. items were not used dangerously (88%). For the item "eats edible items", percent agreement was below criterion, but kappa ratio was 1.00.

B. TEST-RETEST RELIABILITY

Raters agreed about what happens to removed items. Observed percents were 80 to 100. Kappa ratios range from .60 to .78.

TABLE 5 DESCRIPTIVE INFORMATION - REMOVING

		MEAN	STD DEV	MINIMUM	MAXIMUM
INTER-RATER CONDITION					
SEVERITY		58.43	24.41	15	89
FREQUENCY		61.56	41.41	8	150
TEST-RETEST CONDITION					
SEVERITY	T1	51.10	26.67	7	89
	T2	42.45	22.87	2	90
FREQUENCY	T1	69.75	46.23	8	150
	T2	79.90	47.79	10	150

SUMMARY

The question on the Quick Form regarding Removing was answered at a sufficient percent agreement to be considered as having inter-rater and test-retest reliability. In addition, it shows a high level of test-retest reliability as assessed by the kappa ratio. The severity of removing behavior can be answered reliably by raters, though not by pairs. This situation was also seen for the frequency of removing, and for the amount of effort required to remove things. Both inter-rater and test-retest reliability has been demonstrated for the other clinical characteristics of removing.

VERBAL ABUSE

INTER-RATER CONDITION

A total of 147 forms were completed.

Pairs agreed that 45 (10%) residents are verbally abusive and that 361 residents (78%) were not verbally abusive. They disagree about 57 residents (12%). Mean percent agreement was 88% and the kappa ratio was .59.

TEST-RETEST CONDITION

Forty-nine forms were involved. Forty-five (91.8%) were agreements and four (8.1%) were disagreements. The sample's mean percent agreement was 87% and the kappa ratio was .84.

1. HOW SERIOUS A PROBLEM IS THE RESIDENT'S VERBAL ABUSE?

A. INTER-RATER RELIABILITY

When severity was examined as a five point scale, a T of .32, ($p<.001$) one of the highest values of T among the scales, was achieved. When examined as a two point scale, a T of .78 (p,.001) and a kappa ratio of .69 was obtained. For severity, pairs showed the greatest degree of agreement on verbal abuse.

B. TEST-RETEST RELIABILITY

When severity was examined as a five point scale, a T of .32($p<.001$), was achieved. When examined as a two point scale, a T of .78 ($p<.001$) and a kappa ratio of 1.00 was obtained.

2. WHAT IS THE NATURE OF VERBAL ABUSE?

A. INTER-RATER RELIABILITY

Pairs showed agreement when they indicate that residents were verbally abusive by: 1.swearing and cursing, (82%, kappa ratio .77).

B. TEST-RETEST RELIABILITY

Raters showed agreement on all of the categories listed, (except harassment 73%), with agreement from 82 to 98% and kappa ratios from .67 to 1.00.

3. HOW FREQUENTLY IS THE RESIDENT VERBALLY ABUSIVE?

A. INTER-RATER RELIABILITY

Pairs showed agreement regarding frequency, ($T = .32$, $p < .001$). The kappa ratio was .33.

B. TEST-RETEST RELIABILITY

Raters showed agreement regarding frequency, ($T = .47$, $p < .001$, kappa ratio .58).

4. WHOM DOES THE RESIDENT DIRECT VERBAL ABUSE TOWARDS?

A. INTER-RATER RELIABILITY

Pairs showed adequate percent agreement when saying staff (91%) or others (84%) are the recipients of abuse, but not adequate kappa ratios.

B. TEST-RETEST RELIABILITY

Agreement ranged from 80 to 98 percent though only two of the four categories had kappa ratios above .60.

5. WHERE IS THE RESIDENT WHEN BEING VERBALLY ABUSIVE?

A. INTER-RATER RELIABILITY

Pairs agreed sufficiently in their judgements as to whether residents are in private and/or common areas when being abusive to produce adequate percent agreement (73-96%), but not adequate kappa ratios.

B. TEST-RETEST RELIABILITY

The kappa ratio for abuse occurring in common areas was .69, which corresponds to 89% agreement. Abuse occurring in public areas has a 87% agreement but a kappa ratio below .60.

6. WHEN IS THE RESIDENT VERBALLY ABUSIVE?

A. INTER-RATER RELIABILITY

Percent agreement was above criterion for two of the categories, but none of the kappa ratios exceed .6.

B. TEST-RETEST RELIABILITY

Though no category shows adequate percent agreement, one category, treatments, has a kappa ratio above .60.

7. HOW LOUD IS THE VERBAL ABUSE?

A. INTER-RATER RELIABILITY

Pairs agreed about the volume of verbal abuse once it is audible outside the building, (98%, kappa ratio of 1.00).

B. TEST-RETEST RELIABILITY

Raters agreed about the volume of verbal abuse once it is audible throughout a unit, (84%, kappa ratio .75). Prior to that volume, agreement was below criterion.

8. IS IT GENERALLY POSSIBLE TO LIMIT THE VOLUME OF THE RESIDENT'S VERBAL ABUSE SUCH THAT IT IS ONLY AUDIBLE IN ONE ROOM?

A. INTER-RATER RELIABILITY

Agreement is not shown.

B. TEST-RETEST RELIABILITY

Raters do not showed agreement on this question.

TABLE 6 DESCRIPTIVE INFORMATION - VERBAL ABUSE

		MEAN	STD DEV	MINIMUM	MAXIMUM
INTER-RATER CONDITION					
SEVERITY		55.98	24.974	6	150
FREQUENCY		76.67	41.654	2	150
TEST-RETEST CONDITION					
SEVERITY	T1	50.07	22.98	4	91
	T2	40.73	19.74	0	86
FREQUENCY	T1	86.67	42.82	8	150
	T2	101.29	40.14	9	150

SUMMARY

The Verbal Abuse question on the Quick Form shows adequate percent agreement. The kappa ratio for test-retest reliability is acceptable, while that for inter-rater reliability falls just short. Severity and frequency judgments were all adequate. The reliability of items to questions regarding clinical expression of verbal abuse varies, but most questions had a number of reliable response items. This scale is one of the better scales developed in this series.

PHYSICAL ABUSE TOWARD STAFF

INTER-RATER CONDITION

The number of forms completed was 168. This included 106 forms for 54 residents (12%) said to be abusive by both raters, and 62 forms for residents (13%) thought abusive by one rater. The remaining 347 residents (75%) were not abusive.

Mean percent agreement is 87% and the kappa ratio is .58.

TEST-RETEST CONDITION

Thirty-four forms were for 29 agreements (86%) and 5 disagreements (14%). The sample's mean percent agreement is 83% and the kappa ratio is .77.

1. HOW SERIOUS A PROBLEM IS THE RESIDENT'S PHYSICAL ABUSE OF THE STAFF?

A. INTER-RATER RELIABILITY? B. TEST-RETEST RELIABILITY?

As a two-point scale, severity shows high agreement, ($T = .76$, $p < .001$). The kappa ratio was .6.

As a five-point scale, severity shows low but significant agreement, ($T = .22$, $p < .01$).

B. TEST-RETEST RELIABILITY

As a two-point scale, severity shows high agreement, ($T = .79$, $p < .001$). The kappa ratio was 1.00.

As a five-point scale, severity shows no agreement, ($T = .22$, p NS).

2. IN WHAT WAY IS THE RESIDENT PHYSICALLY ABUSIVE TO STAFF?

A. INTER-RATER RELIABILITY

Pairs were close to showing agreement on the most severe form of abuse, but did not achieve it. Kappa ratios were below the criterion.

B. TEST-RETEST RELIABILITY

Raters showed agreement regarding the severe form of abuse, (86%, kappa ratio of .74) which occurs infrequently, but appear not to distinguish between the two lessor forms of abuse.

3. HOW FREQUENTLY IS THE RESIDENT ABUSIVE?

A. INTER-RATER RELIABILITY

Pairs yield a significant level of agreement concerning the frequency of abuse, ($T = .24$, $p < .01$). The kappa ratio was .26.

B. TEST-RETEST RELIABILITY

Raters also yield a significant level of agreement concerning the frequency of abuse, ($T = .36$, $p < .01$). The kappa ratio was .39.

4. HAS THE RESIDENT CAUSED A WCB INJURY TO A STAFF MEMBER WITHIN THE LAST YEAR?

A. INTER-RATER RELIABILITY

Nurses are not necessarily informed when a WCB claim is made. Responses to this question were "yes" and "not to my knowledge". Agreement 87%. The kappa ratio was .43.

B. TEST-RETEST RELIABILITY

Total agreement was seen, with 100% and a kappa ratio of 1.00.

5. WHAT IS THE RESIDENT'S SIZE?

HEIGHT

A. INTER-RATER RELIABILITY

While 73% agreement was seen, the kappa ratio was .63.

B. TEST-RETEST RELIABILITY

While 76% agreement was seen, the kappa ratio was again .63.

WEIGHT

A. INTER-RATER RELIABILITY

Inter-rater agreement was not shown for resident weight, (67%, kappa ratio .45).

B. TEST-RETEST RELIABILITY

Kappa ratio for test-retest agreement was .77, though the percent figure was 76%.

6. WHAT ARE THE RESIDENT'S PHYSICAL ABILITIES?

STRENGTH

A. INTER-RATER RELIABILITY

Inter-rater agreement was shown for resident strength by the kappa ratio of .63, but agreement was 67%.

B. TEST-RETEST RELIABILITY

Test-retest agreement was .35 as assessed by the kappa ratio, a level which was below criterion, as was the percent of 55.

6. MOBILITY

A. INTER-RATER RELIABILITY

Agreement on the four categories of mobility ranged from 74-100%, with two kappa ratios above criterion.

B. TEST-RETEST RELIABILITY

Agreement ranged from 86 to 97%. The kappa ratios were 1.00 for 2 categories.

TABLE 7 DESCRIPTIVE INFORMATION - PHYSICAL ABUSE TO STAFF

		MEAN	STD DEV	MINIMUM	MAXIMUM
INTER-RATER CONDITION					
SEVERITY		52.32	23.52	2	129
FREQUENCY		74.77	41.98	1	150
HEIGHT		54.56	37.29	4	130
WEIGHT		48.35	33.98	2	128
STRENGTH		71.60	33.59	5	125
TEST-RETEST CONDITION					
SEVERITY	T1	55.03	22.11	9	100
	T2	47.28	28.35	3	143
FREQUENCY	T1	90.28	39.27	14	150
	T2	91.97	44.88	9	150
HEIGHT	T1	57.45	35.12	5	123
	T2	57.83	36.65	5	124
WEIGHT	T1	52.76	33.78	5	122
	T2	57.35	32.95	6	123
STRENGTH	T1	74.31	35.17	5	124
	T2	68.38	33.74	26	129

SUMMARY

The Quick Form questions regarding physical abuse of staff indicate acceptable, or nearly acceptable levels, of inter-rater and test-retest reliability. Severity can be rated on a two-point or five-point ordinal scale. Frequency can also be rated acceptably. Raters, though not pairs, were reliable in their indications of forms of abuse, and knowledge of WCB claims. Judgements regarding resident height, weight and strength had the curious result of showing percent agreement below 80%, yet kappa ratios above .6. One would think that concrete characteristics such as these would be easy to rate, and it is curious that nurses found them of moderate difficulty.

PHYSICAL ABUSE OF OTHER RESIDENTS

INTER-RATER CONDITION

Fifty forms were used, 28 for 14 residents (3.0%) rated by both nurses as abusing other residents. One member of the rating pair claimed an additional 22 residents (4.8%) to be abusive while 427 residents (92.0%) were rated as not abusive.

Mean percent agreement is 95% and kappa ratio is .64.

TEST-RETEST CONDITION

Twenty-five forms were completed for 23 agreements (92%) and 2 disagreements (8%). The sample's mean percent agreement was 92% and the kappa ratio was .89.

1. HOW SERIOUS IS THE RESIDENT'S ABUSE OF OTHER RESIDENTS?

A. INTER-RATER RELIABILITY

Reliability was quite high ($T = .71$, $p < .01$; kappa ratio 1.00) for a two-point scale, but not for a five-point scale ($T = -.14$, p NS).

B. TEST-RETEST RELIABILITY

Reliability was quite good for both two-point scales ($T = .71$, $p < .01$; kappa ratio 1.00) and five point scale ($T = .30$, $p < .02$).

2. HOW FREQUENTLY HAS THE RESIDENT ACTUALLY ABUSED ANOTHER RESIDENT?

A. INTER-RATER RELIABILITY

Frequency shows good agreement, ($T = .33$, $p < .05$), but not an acceptable kappa ratio.

B. TEST-RETEST RELIABILITY

Frequency shows high agreement, ($T = .42$, $p < .001$), but again the kappa ratio was below .6.

3. WHO IS USUALLY RESPONSIBLE WHEN THIS RESIDENT IS INVOLVED IN AN ALTERCATION?

A. INTER-RATER RELIABILITY

Pairs agreed totally (100%) that the recipient was never responsible for abuse received, while also having quite high agreement (86%) that sometimes they can not say who was responsible. Results for the items of shared responsibility and complete guilt by the rated resident were both below criterion.

B. TEST-RETEST RELIABILITY

Agreement ranged from 74 to 100% while kappa ratios were above .6 for two of the four categories.

4. HOW SERIOUS IS THE NATURE OF THE PHYSICAL ABUSE?

A. INTER-RATER RELIABILITY

Pairs agreed completely that potentially lethal forms of abuse do not occur. They did not appear to distinguish lesser forms of abuse.

B. TEST-RETEST RELIABILITY

Raters agreed completely that potentially lethal forms of abuse do not occur. The two lessor forms of abuse had percent agreements above 80%; one shows a kappa ratio above .60, and the other was .57.

5. WHAT IS THE CONSEQUENCE OF THE ABUSE FOR THE RECIPIENT?

A. INTER-RATER RELIABILITY

Pairs showed total agreement that medical or hospital care was never required because of resident abuse. Good agreement was seen regarding physical evidence (86%).

B. TEST-RETEST RELIABILITY

Raters showed agreement ranging from 87 to 100%, with kappa ratios of 1.00.

TABLE 8 DESCRIPTIVE INFORMATION - PHYSICAL ABUSE OF OTHER RESIDENTS

		MEAN	STD DEV	MINIMUM	MAXIMUM
INTER-RATER CONDITION					
SEVERITY		57.29	29.50	9	127
FREQUENCY		107.00	43.87	8	150
TEST-RETEST CONDITION					
SEVERITY	T1	48.65	29.48	7	124
	T2	44.70	27.14	0	123
FREQUENCY	T1	117.09	41.11	9	150
	T2	123.13	40.40	8	150

SUMMARY

The question from the Quick Form regarding physical abuse of other residents shows good inter-rater and test-retest reliability. Severity shows inter-rater reliability as a two-point scale and test-retest reliability as a five-point scale. Frequency showed good reliability, both for inter-rater and test-retest situations. The abuse engaged in was not apt to cause the need for medical attention, but does leave physical evidence. The scale is acceptable for use.

SEXUAL MISBEHAVIORS

INTER-RATER CONDITION

Five residents (1.1%) were agreed as showing sexual misbehaviors, and 9 (1.9%) were in dispute, for a total of 19 forms. The remaining 449 (97.0%) were agreed as not engaging in sexual misbehaviors. Ninety-eight percent agreement and a .71 kappa ratio were found.

TEST-RETEST CONDITION

Of the eight forms, two (25%) were agreements and six (75%) were disagreements. The agreement was 95%, and the kappa ratio was 1.00!

1. HOW SERIOUS IS THE RESIDENT'S SEXUAL MISBEHAVIOR(S)?

A. INTER-RATER RELIABILITY

Pairs agreed on the severity categories all five residents were placed in.

B. TEST-RETEST RELIABILITY

Raters placed one resident in the same severity category and one resident in a different category.

2. HOW FREQUENTLY DOES THE RESIDENT MAKE UNACCEPTABLE SEXUAL COMMENTS TO OTHER RESIDENTS?

A. INTER-RATER RELIABILITY

Pairs agreed completely that none of the residents make such comments to other residents.

B. TEST-RETEST RELIABILITY

One resident was rated as not making such comments at both Time One and Time Two. The other resident was said to make such comments at

both rating times, but the frequency differed between the two occasions.

HOW FREQUENTLY DOES THE RESIDENT MAKE UNACCEPTABLE SEXUAL COMMENTS TO STAFF MEMBERS?

A. INTER-RATER RELIABILITY

Agreement was seen for only one resident, (20%).

B. TEST-RETEST RELIABILITY

Both residents were rated as making sexual comments to staff members on both occasions, but the frequency differed.

3. DOES THE RESIDENT TRY TO "GET FRESH" WITH STAFF MEMBERS?

A. INTER-RATER RELIABILITY

Agreement was seen for two residents, (40%).

B. TEST-RETEST RELIABILITY

One resident was rated as "no" on both occasions. The other was rated as "yes" Time One and "no" Time Two.

4. DOES THE RESIDENT TRY TO TOUCH OTHER RESIDENTS IN UNWELCOMED AND INAPPROPRIATE WAYS?

A. INTER-RATER RELIABILITY

Pairs agreed (100%) that no resident was currently touching other residents.

B. TEST-RETEST RELIABILITY

Raters agreed again that neither resident was touching other residents.

5. DOES THE RESIDENT EXPOSE HIS/HER BODY IN INAPPROPRIATE PLACES?

A. INTER-RATER RELIABILITY

Pairs agreed about three residents (60%) they said were not exposing themselves, but disagreed about the other two.

B. TEST-RETEST RELIABILITY

Neither resident was said to expose their body on either occasion.

6. DOES THE RESIDENT MASTURBATE OPENLY?

A. INTER-RATER RELIABILITY

Pairs agreed (100%) that no resident was masturbating openly.

B. TEST-RETEST RELIABILITY

Responses to these questions were "no" on both occasions for both residents.

TABLE 9 DESCRIPTIVE INFORMATION - SEXUAL MISBEHAVIORS

	MEAN	STD DEV	MINIMUM	MAXIMUM
SEVERITY	61.80	18.63	37	104
COMRES	10.30	32.57	0	103
COMMEM	56.90	52.91	0	142
FRESH	27.60	45.11	0	135
TOUCH	.00	.00	0	0
EXPOSE	28.60	47.55	0	122
MAS	.00	.00	0	0

SUMMARY

The small number of residents engaged in some form of sexual misbehavior makes evaluation of this scale tentative, but it does appear from the available data to be reliable.

MEALTIME BEHAVIORS

INTER-RATER CONDITION

Sixty-three residents were rated as having mealtime misbehaviors by two nurses, 113 by one nurse, and 287 by neither nurse. Mean percent agreement was 76% and kappa ratio was .43.

TEST-RETEST CONDITION

A total of 43 forms were completed, 32 agreements (74%) and 11 disagreements (26%). Mean percent agreement was 82% and kappa ratio was .60.

1. HOW SEVERE ARE MEALTIME MISBEHAVIORS?

A. INTER-RATER RELIABILITY

As a two point-scale, agreement was quite high, ($T = .59$, $p < .001$). The kappa ratio was below .6. As a five-point scale, no agreement was seen, ($T = .10$, p NS).

B. TEST-RETEST RELIABILITY

As a two point-scale, agreement was quite high, ($T = .56$, $p < .01$), though the kappa ratio was below .6. For a five-point scale, no agreement was seen, ($T = .08$, p NS).

2. MARK SPECIFIC DIFFICULTIES

A list of 13 specific mealtime problems was included, and can be checked in Appendix One.

A. INTER-RATER RELIABILITY

Agreement on these questions was generally quite high with nine above 80% and eight kappa ratios above .6.

B. TEST-RETEST RELIABILITY

Agreement was again quite high with nine of the 13 categories above 80 percent and eight kappa ratios above .60.

TABLE 10 DESCRIPTIVE INFORMATION - MEALS BEHAVIORS

		MEAN	STD DEV	MINIMUM	MAXIMUM
INTER-RATER CONDITION					
SEVERITY		56.00	23.26	12	145
TEST-RETEST CONDITION					
SEVERITY	T1	59.63	23.86	3	92
	T2	49.84	19.76	0	89

SUMMARY

The question from the Quick Form regarding mealtime misbehaviors just misses reaching an acceptable level of inter-rater reliability, but achieves sufficient test-retest reliability. Severity was well rated as an two-point ordinal scale. Specific difficulties were largely reliable.

BED

QUICK FORM QUESTION

DOES THE RESIDENT TRY TO CLIMB OUT OF BED EVEN THOUGH HE/SHE LACKS THE ABILITY TO DO SO SAFELY?

INTER-RATER CONDITION

Forty-eight questions were answered "yes", for 18 residents (4%) rated by both raters and 30 (6%) by one rater. The remaining 415 residents (90%) were rated "no" by both raters.

The mean percent agreement was 94% and the kappa ratio was .53.

TEST-RETEST CONDITION

Twenty-six "yes" responses included 20 (4%) agreements and 6 (1%) disagreements. Agreement was 99% and the kappa ratio was .90.

SPECIAL SCALE QUESTION

1. IS IT TYPICAL OF THIS RESIDENT TO TRY TO CLIMB OUT OF BED?

A. INTER-RATER RELIABILITY

Of the 18 residents said to try to climb out of bed, it was said to be typical of 15 (83%).

B. TEST-RETEST RELIABILITY

Of the 20 residents, trying to climb out of bed was said to be typical of 15 (75%).

WHAT IS THE RESULT?

A. INTER-RATER RELIABILITY

Four choices were offered (will fall, may fall, able with assistance, fully able). Only one has an agreement above 80%, and, none of the three choices which allow calculation of a kappa exceed .6

B. TEST-RETEST RELIABILITY

Two choices showed agreement above 80%, but none shows an acceptable kappa ratio.

QUICK FORM QUESTION

DOES THE RESIDENT RESIST BED RESTRAINTS?

INTER-RATER CONDITION

Sixteen forms were completed including two residents (<1%) said by both raters to resist restraints and 14 residents (3%) said to do so by one rater. Mean agreement was 97% and kappa ratio was .23.

TEST-RETEST CONDITION

Twelve forms were completed, including two agreements (17%) and 10 disagreements (83%).

SPECIAL SCALE QUESTION

2. BED RESTRAINT: WHICH ONE OF THE BELOW IS MOST CORRECT?

is not necessary

is not used because the resident is so resistive he or she might be injured

is used and mildly resisted

is used and tolerated by the resident

A. INTER-RATER RELIABILITY

No agreement was shown on items regarding bed restraints. They were used quite infrequently by the time ratings were completed, but were more prevalent when the PBAT was being developed.

B. TEST-RETEST RELIABILITY

Eighty to 90 percent agreement was seen here, with 3 of the 4 categories having kappa ratios above .60.

SUMMARY

This scale was unreliable and currently irrelevant and should be discarded.

WANDERING

INTER-RATER CONDITION

The number of forms completed was 47.

Pairs considered ten residents (2.2%) to be wanderers, disagreed about 27 residents (2.2%) and eliminated 426 residents (92.0%). Mean percent agreement was 94% and kappa ratio was .56.

TEST-RETEST CONDITION

Nineteen forms were completed, 17 agreements (89%) and 2 disagreements (11%). The mean percent agreement was 89% and the kappa ratio was .86.

1. HOW SERIOUS IS THE RESIDENT'S WANDERING?

A. INTER-RATER RELIABILITY

The severity scale was not meaningful to pairs. Severity does not approach significance or reasonable agreement as either a two-point scale, ($T = .20$, p NS) or as a five-point scale, ($T = -.20$, p NS).

B. TEST-RETEST RELIABILITY

The severity scale was not meaningful to raters. Severity does not approach significance or reasonable agreement as either a two-point scale, ($T = .44$, p NS) or as a five-point scale, ($T = .19$, p NS).

2. HOW FREQUENTLY DOES THE RESIDENT WANDER INTO OTHER RESIDENT'S ROOMS?

A. INTER-RATER RELIABILITY

Pairs agreed quite well on the frequency of wandering into other residents rooms, (kappa ratio was .72).

B. TEST-RETEST RELIABILITY

Raters did not showed agreement on this question, (kappa ratio was .43).

HOW FREQUENTLY DOES THE RESIDENT WANDER IN OTHER AREAS ON THE UNIT?

A. INTER-RATER RELIABILITY

Kappa ratio was .75, indicating good inter-rater agreement.

B. TEST-RETEST RELIABILITY

Raters do agreed on the frequency of wandering into unit areas, (kappa ratio .77).

HOW FREQUENTLY DOES THE RESIDENT WANDER OFF THE UNIT BUT IN THE BUILDING?

A. INTER-RATER RELIABILITY

Inter-rater agreement was 1.00 according to the kappa ratio.

B. TEST-RETEST RELIABILITY

Kappa ratio was .64, indicating test-retest agreement on this question.

HOW FREQUENTLY DOES THE RESIDENT WANDER OUTSIDE THE BUILDING?

A. INTER-RATER RELIABILITY

Kappa ratio is .66. Percent agreement was 70%.

B. TEST-RETEST RELIABILITY

No agreement was seen, (kappa ratio was .41).

3. ON WHICH SHIFTS DOES THE RESIDENT WANDER?

A. INTER-RATER RELIABILITY

Pairs showed a high degree of agreement regarding the time of day when residents wander, (90-100%, kappa ratios criterion).

B. TEST-RETEST RELIABILITY

Raters showed a high degree of agreement regarding the time of day when residents wander, with 89 to 100% agreement and all three kappa ratios equal to 1.00.

4. HOW DIFFICULT IS IT TO PREVENT OR LIMIT THIS RESIDENT'S WANDERING?

A. INTER-RATER RELIABILITY

Possible choices included less difficult, similar to, and more difficult than most wanderers. Pairs were best able to agree about wanderers who were either less difficult (80%) or more difficult to control, (90%). Two kappa ratios were 1.00, and one can not be calculated.

B. TEST-RETEST RELIABILITY

Percent agreement was 78 to 89% and all kappa ratios exceed .60.

5. ONCE THE RESIDENT HAS WANDERED, WHAT IS HIS RESPONSE TO STAFF ATTEMPTS TO RETURN HIM TO A PROPER AREA?

A. INTER-RATER RELIABILITY

Of the four choices, (cooperates, objects, resists, becomes aggressive) pairs were best able to agree about the two extremes. When choosing "cooperates with return" pairs showed high agreement, (90%, kappa ratio of 1.00), as they also did when choosing "physically aggressive to return", (90%, kappa ratio 1.00).

B. TEST-RETEST RELIABILITY

When choosing "physically aggressive to return", 83 percent agreement was seen with a kappa ratio of 1.00. The other choices did not meet the criterion.

6. IS THE RESIDENT WANDERING IN A WHEELCHAIR, WALKER OR BY WALKING?

A. INTER-RATER RELIABILITY

Pairs showed very good agreement regarding the mode of transport used by wanderers. Agreement ranged from 90-100%, and kappa ratio was either 1.00 or not calculable.

B. TEST-RETEST RELIABILITY

Raters showed complete (100%) agreement regarding the mode of transport used by wanderers. Kappa ratio was 1.00 or no calculable.

TABLE 11 KAPPA RATIOS FOR FREQUENCY OF WANDERING

PLACE	AGREE	TOTAL	KAPPA	KAPPA MAX	KAPPA RATIO
other resident's rooms	9	18	.33	.78	.43
areas on the unit	10	18	.42	.64	.66
off the unit	9	18	.37	.58	.64
outside the building	11	18	.23	.56	.41

TABLE 12 DESCRIPTIVE INFORMATION - WANDERING

		MEAN	STD DEV	MINIMUM	MAXIMUM
INTER-RATER CONDITION					
SEVERITY		61.60	22.79	33	95
ROOM		80.40	59.28	0	150
AREAS		63.80	52.43	0	150
INBLDG		52.00	58.74	0	150
OUTBLDG		119.50	51.98	0	150
TEST-RETEST CONDITION					
SEVERITY	T1	56.72	27.12	12	95
	T2	49.11	19.15	2	76
ROOM	T1	80.17	47.28	11	150
	T2	79.00	51.42	7	150
AREAS	T1	75.44	48.50	11	150
	T2	76.22	48.89	8	150
IN BLDG	T1	44.28	51.77	0	150
	T2	88.17	57.19	0	150
OUT BLDG	T1	95.72	63.85	0	150
	T2	110.67	61.20	0	150

SUMMARY

The Quick Form question regarding wandering has good inter-rater percent agreement and just misses an acceptable kappa ratio. Test-retest reliability was quite good. While severity can not be meaningfully judged, most other clinical characteristics were well handled.

NOISE

INTER-RATER CONDITION

A total of 164 forms included 110 for 55 residents (12%) named as noisy by both raters and 54 (12%) were disputed. The remaining 354 residents (76%) were not indicated as being noisy. Mean agreement was 88% and kappa ratio was .62.

TEST-RETEST CONDITION

Forty-nine forms were completed, of which 46 (94%) were agreements and 3 were disagreements (6%). Agreement was 91% with a kappa ratio of .89.

1. HOW SERIOUS A PROBLEM IS UNNECESSARY NOISE CREATED BY THIS RESIDENT?

A. INTER-RATER RELIABILITY

For a two-point scale, agreement was very high, ($T = .64$, $p < .001$). For a five-point scale, pairs showed significant agreement about the severity of noise, ($T = .20$, $p < .02$).

B. TEST-RETEST RELIABILITY

Similar to the above, for a two-point scale, agreement was very high, ($T = .64$, $p < .001$). For a five-point scale, raters had a significant agreement about the severity of noise, ($T = .36$, $p < .001$).

2. HOW FREQUENTLY DOES THIS RESIDENT CREATE DISRUPTIVE AND UNNECESSARY NOISE?

A. INTER-RATER RELIABILITY

Pairs demonstrate significant agreement about the frequency of noise, ($T = .25$, $p < .01$).

B. TEST-RETEST RELIABILITY

Raters did not showed agreement about the frequency of noise, ($T = .10$, NS). The kappa ratio was .07.

3. ON WHAT SHIFTS DOES THIS OCCUR?

A. INTER-RATER RELIABILITY

Agreement for the time of day when noise occurs was good. Days (85%, with a kappa ratio of .73) and evenings (87%) exceed criterion.

B. TEST-RETEST RELIABILITY

Agreement was very high, with from 89 to 96% agreement and kappa ratios from .81 to 1.00.

4. WHAT IS THE NATURE OF THE NOISE AND HOW LOUD IS IT?

A. INTER-RATER RELIABILITY

Pairs can handle the matrix choice system this question presents quite well. Seven of the percent agreements exceed the criteria of 80%, as did two kappa ratios.

B. TEST-RETEST RELIABILITY

Eight of the percent agreements exceed criterion.

5. WHEN THE RESIDENT IS BEING NOISY, IS THE NOISE CONTINUOUS OR INTERMITTENT?

A. INTER-RATER RELIABILITY

Pairs did not agree on these choices.

B. TEST-RETEST RELIABILITY

Raters indicate agreement on continuous noise (80%, kappa ratio of .69) and were close to criteria for intermittent noise.

6. WHEN THE RESIDENT IS BEING NOISY, ARE YOU ABLE TO QUIET THEM?

A. INTER-RATER RELIABILITY

No agreement was seen.

B. TEST-RETEST RELIABILITY

Adequate percent agreement was shown for one choice and acceptable kappa ratios for all three.

TABLE 13 DESCRIPTIVE INFORMATION - NOISE

		MEAN	STD DEV	MINIMUM	MAXIMUM
INTER-RATER CONDITION					
SEVERITY		60.25	24.05	7	145
FREQUENCY		58.35	41.25	1	150
TEST-RETEST CONDITION					
SEVERITY	T1	57.41	21.39	5	115
	T2	55.89	22.32	21	140
FREQUENCY	T1	70.15	37.26	7	140
	T2	75.17	41.99	4	150

SUMMARY

The question of the Quick Form leading to this scale was reliable, both in inter-rater and test-retest forms. Severity can be judged on a five-point ordinal scale and frequency can be rated between judges.

Other clinical characteristics were almost all rated reliably. This scale is appropriate for practical use.

SEVERITY

Since most of the scales require a severity rating, a question one might ask is whether rater pairs agreed about the severity of problem behaviors they observe. What is the nature of rater's use of the severity measure? As presented on the PBAT sheets, severity has the superficial appearance of a continuous variable with five verbal labels attached along the way (see Appendix 2). Are raters using the severity measure as a continuous variable, or as a variable with discrete categories? Several analyses were done to examine this issue.

First, severity can be considered a continuous variable. A simple Pearson correlation (raw centimeter score) of Rater A and Rater B's judgments was done, and is presented in Tables 14 and 15. As can be seen from the table, the median correlation is approximately .5 for the inter-rater situation and .6 for the test-retest situation. These results were not as high as one might wish. If the severity measure is a continuous variable, it does not allow a close correlation between rater's responses. A closer look at the rater's use of the severity measure is warranted.

Second, the severity item can be broken into discrete categories according to the five verbal labels along the line. T values for severity are given in Table 16. In assessing inter-rater reliability, three (Verbal Abuse, Physical Abuse to Staff and Noise) of the nine scales which require severity ratings showed low but significant T values, indicating ability to use severity as a five point scale only on those specific scales. In assessing test-retest

reliability, again, three of the nine scales (Verbal Abuse, Physical Abuse to Residents, and Noise) had significant T values.

Third, severity can be broken into two categories, moderate and severe. Results are shown in Table 17. On inter-rater reliability, five of the nine scales had significant T values, and four scales had kappa ratios above .60. In the test-retest condition, six had significant T values and five had kappa ratios above .6. These results suggest that the severity measure can be used as a variable with either five categories or two categories depending on the scale.

TABLE 14 CORRELATION OF RATER PAIR'S SEVERITY AND FREQUENCY MEASURES - INTER-RATER CONDITION

RATER	SEVERITY WITH SEVERITY	FREQUENCY WITH FREQUENCY	SEVERITY WITH FREQUENCY
01	.18	.84	-.75
02			-.67
03	-.14	.98	-.87
04			.03
05	.23	.86	-.54
06			-.38
09	-.47	.32	-.32
07			-.17
08	.72	.49	-.22
10			-.42
11	.62	.66	-.23
13			.18
12	.20	.48	-.21
13			-.78
14	.36	.80	-.60
15			-.26
16	.57	.68	-.81
17			-.23
18	.14	.65	-.26
19			-.66
20	.28	.36	.02
21			-.23
22	.65	.37	-.62
23			-.14
22	.65	.54	-.07
24			-.21
25	.80	.18	-.59
28			-.72
26	.76	~	~
27			~

~ values cannot be calculated for this pair

TABLE 15 CORRELATION OF RATER'S SEVERITY AND FREQUENCY MEASURES TEST-RETEST CONDITION

RATER	SEVERITY WITH SEVERITY	FREQUENCY WITH FREQUENCY	SEVERITY WITH FREQUENCY
01	.75	.93	- .87
02	.06	.77	- .85
03	--	.--	.--
04	.51	- .50	- .99
05	- .05	.31	- .83
06	.27	.39	- .90
09	1.00	-1.00	1.00
07	.18	.42	- .82
08	.63	.88	- .43
10	.66	.91	- .49
11	.64	.77	- .43
13	.38	.50	- .54
12	- .13	.80	- .67
13	.38	.50	- .54
14	.64	.89	- .53
15	- .01	.63	- .65
16	.50	.99	-1.00
17	.--	.--	.--
18	.07	- .25	- .75
19	.56	.85	- .46
20	1.00	1.00	-1.00
21	.90	.57	- .91
22	.79	.77	- .41
23	.61	.60	- .09
22	.79	.77	- .41
24	.81	.77	- .15
25	.89	.89	- .71
28	.68	.09	- .82
26	.61	.97	- .89
27	--	--	--

-- values cannot be calculated for these raters

TABLE 16 CAN SEVERITY BE USED AS A FIVE POINT SCALE?

		N	AGREE	DISAGREE	X2	P	T
Rummaging	T1	6	2	4	.22		.11
	T2	10	2	8	.40		-.07
Removing	T1	8	1	7	1.17		-.17
	T2	20	9	11	3.80		.27
Verbal Abuse	T1	45	22	23	13.09	.001	.32
	T2	45	22	23	13.09	.001	.32
Physical Abuse							
To staff	T1	51	21	30	6.71	.01	.22
	T2	29	12	17	3.76		.22
To residents	T1	14	2	12	1.24		-.14
	T2	23	11	12	5.84	.02	.30
Sex Misbehaviors	T1	5	3	2	2.60		.47
	T2	2	1	1	.67		.33
Meals	T1	64	21	43	1.90		.10
	T2	32	10	22	.54		.08
Wandering	T1	10	1	9	1.73		-.20
	T2	18	7	11	1.56		.19
Noise	T1	55	22	33	6.22	.02	.20
	T2	46	24	22	17.42	.001	.36

X2	P
3.84	.05
5.41	.02
6.64	.01
10.83	.001

TABLE 17 CAN SEVERITY BE USED AS A TWO POINT SCALE?

ID	RA	RB	RC	RD	S	AP	NP	DP	MP	X2	P	T	K	KR
21	3	1	1	1	6	50	33	17	67	.83		.33	.25	.25
31	6	0	4	0	10	60	40	0	60	.50		.20	.00	.00
22	5	0	2	1	8	63	25	13	75	2.13		.50	.38	1.00
32	16	1	1	2	20	80	10	10	90	12.85	.001	.80	.61	.61
23	34	2	3	6	45	76	11	13	89	27.24	.001	.78	.64	.69
33	39	0	5	1	45	87	11	2	89	27.24	.001	.78	.26	1.00
24	41	4	2	4	51	80	12	8	88	29.84	.001	.76	.50	.60
34	24	0	3	2	29	83	10	7	90	18.28	.001	.79	.52	1.00
25	10	2	0	2	14	71	14	14	86	7.21	.01	.71	.59	1.00
35	20	0	1	2	23	87	4	9	96	19.22	.001	.91	.78	1.00
26	4	1	0	0	5	80	20	0	80	2.00		.60	.00	.00
36	1	0	1	0	2	50	50	0	50	.50		.00	.00	.
27	48	2	11	3	64	75	20	5	80	22.58	.001	.59	.23	.49
37	24	1	6	1	32	75	22	3	78	10.16	.01	.56	.14	.36
29	5	1	3	1	10	50	40	10	60	.50		.20	.09	.17
39	13	0	5	0	18	72	28	0	72	3.61		.44	.00	.00
20	38	6	4	7	55	69	18	13	82	22.29	.001	.64	.47	.52
30	37	2	3	4	46	80	11	9	89	28.20	.001	.78	.55	.61

LEGEND
ID = SCALE
RA, RB, RC, RD
S= TOTAL N
AP= NP= DP= MP=
X2= T= P=
K= Kappa KMAX= Kappa maximum KR= Kappa ratio

FREQUENCY

Another characteristic of problem behavior rated on many of the Special Scales is frequency of occurrence. While severity is an abstract feature that quite easily can be thought of as affected by the rater's attitude and personality, frequency is a more objective characteristic. Do the raters showed greater agreement on frequency than on severity?

First, frequency can be examined as a continuous variable. Tables 14 and 15 show inter-rater and test-retest Pearson correlations. The median Pearson correlation is approximately .6 in the inter-rater condition and .7 for the test-retest condition. These correlations are slightly higher than those for severity.

Second, frequency can be thought of as an ordinal scale variable with four categories. T values were significant in all but two instances, as shown on Table 18.

Frequency appears to be an easier variable for raters to use than severity.

TABLE 18 FREQUENCY

SCALE	TIME	A	D	N	T VALUE	P
RUMMAGING	T1	4	2	6	.56	.05
	T2	8	2	10	.73	.001
REMOVING	T1	4	4	8	.33	NS
	T2	12	8	20	.47	.001
VERBAL ABUSE	T1	22	23	45	.32	.001
	T2	27	18	45	.47	.001
PHYSICAL ABUSE TO STAFF	T1	22	29	51	.24	.01
	T2	15	14	29	.36	.01
TO RESIDENTS	T1	7	7	14	.33	.05
	T2	13	10	23	.42	.001
NOISE	T1	24	31	55	.25	.01
	T2	15	16	46	.10	NS

LEGEND
A= AGREE
D= DISAGREE
N= TOTAL NUMBER OF RATINGS

OTHER LINE BISECTIONS

Two scales had items in addition to severity and frequency which are answered by bisecting lines. On the Rummaging scale, the item questions the rater's ability to stop a resident from rummaging. On the Physical Abuse to Staff scale, items regarding height, weight and strength are all answered by bisecting a line. Statistics for these items are given in Table 19.

The items can be thought of as ordinal scale variables of three categories. T values for stop rummaging were not significant. T values for all the physical characteristics were significant and indicate a moderate level of agreement.

These results indicate that while the stop rummaging question can not reasonably be used by nurses, the physical characteristics

questions can be used well. The line bisection method of indicating a response to questions can be used appropriately.

TABLE 19 OTHER LINE BISECTIONS

SCALE	TIME	A	%	D	TOTAL	T	P
RUMMAGING							
STOP	T1	1	17	5	6	*	*
	T2	6	60	4	10	.40	NS
PHYSICAL ABUSE TO STAFF							
HEIGHT	T1	37	73	14	51	.59	.001
	T2	22	76	7	29	.64	.001
WEIGHT	T1	34	67	17	51	.50	.001
	T2	22	76	7	29	.64	.001
STRENGTH	T1	34	67	17	51	.50	.001
	T2	16	55	13	29	.33	.02

LEGEND
A= number of agreements
%= percent agreement
D= number of disagreements
T= T value
P= probability
* T should not be calculated if the number of agreements is less than expected by chance, i.e., with three categories, chance is .33; with six judgements, two could be significant by chance yet only one agreement is observed.

A PROPOSED SECOND VERSION OF A PROBLEM BEHAVIOR ASSESSMENT INSTRUMENT

The preceding section documents results obtained from an investigation of the reliability of the PBAT. These results could be used to devise a second version of a problem behavior assessment instrument. Such an instrument could also incorporate a few features which experience with the PBAT suggested to the author. The Problem Behavior Assessment Instrument (PBAI) is given in Appendix Four. The PBAI retains the more successful elements of the PBAT: (a) the Quick Form (renamed the Survey Form, after minor changes) and (b) seven of the Special Scales (Rummaging, Removing, Verbal Abuse, Physical Abuse to Staff, Physical Abuse to Other Residents, Wandering and Noise). No new items are added. Items regarding frequency have replaced verbal anchors with numerical anchors, which requires caution in interpreting results obtained in this manner.

COMPARISON OF RATERS AND RATER PAIRS

Many reliability studies involve only one set of raters, or only a few raters. The present study, because of the nature of the site, has 15 pairs of raters, from a group of 28 raters. It is as though 15 inter-rater replications were done, and 28 test-retest replications. Previous sections have examined the reliability of questions across pairs or individual raters. The following sections allow comparison of results among rater pairs and among individual raters, across questions. Tables 20 to 23 summarize this information.

INTER-RATER RELIABILITY OF RATER PAIRS

On the First Part of the Quick Form, rater pairs had an agreement percent range of 70 to 95%. Only two pairs were below 80%. Eight of

the 15 pairs had kappa ratios below .60, with the other ratios going as high as .93.

On the Second Part of the Quick Form, percent agreement ranged from 85 to 98%. Only one pair has a kappa ratio below .60, while the highest kappa ratio was .83.

TEST-RETEST RELIABILITY OF RATERS THEMSELVES

On the First Part of the Quick Form, percent agreement ranged from 70 to 100%, with eight of the 28 raters having values below 80%. Three of the raters had effective percents above 80%. Kappa ratios were below .60 for five raters, while 16 raters had kappa ratios of 1.00!

On the Second Part of the Quick Form, agreement ranged from 80 to 95%. Again, three of the raters had effective percents above 80%. Kappa ratios were below .6 for only 1 of the 28 raters, while eight raters had kappa ratios equal to 1.00!

SUMMARY

This section highlighted the diversity of results seen among pairs and among individual raters. In both rating conditions, more pairs and raters met criterion for the second more objective section of the Quick Form than for the first more attitudinal section. It showed that nurses vary in their ability to use the PBAT in its current state, and suggests that they may also vary in ability to rate behavior generally. One can surmise that with additional training or preparation for raters, the PBAT could yield more uniformly reliable results.

TABLE 20 INTER-RATER RELIABILITY OF RATER PAIRS

FIRST PART OF THE QUICK FORM

PAIR	YY	%	NY YN	%	NN	%	EP %	DP %	MP %	K	K MAX	K RATIO
1	10	5	27	12	183	83	27	61	88	.36	.88	.41
2	8	3	24	10	208	87	25	65	90	.35	.89	.39
3	14	6	11	5	195	89	56	39	95	.69	.75	.93
4	33	10	54	16	253	74	38	46	84	.45	.88	.52
5	31	9	102	30	207	61	23	47	70	.26	.30	.86
6	22	7	94	28	214	65	19	53	72	.23	.26	.87
7	29	8	71	19	270	73	29	52	81	.35	.57	.62
8	13	7	33	18	134	74	28	53	82	.34	.78	.43
9	9	4	23	11	168	84	28	60	89	.38	.92	.41
10	50	14	58	17	242	69	46	37	83	.53	.69	.77
11	35	9	67	18	278	73	34	48	82	.41	.81	.50
12	56	16	52	14	252	70	52	34	86	.59	.77	.78
13	60	16	68	18	242	65	47	35	82	.52	.82	.64
14	18	5	77	20	285	75	19	61	80	.23	.47	.49
15	12	3	19	5	319	91	39	56	95	.53	.93	.57

TABLE 21 INTER-RATER RELIABILITY OF RATER PAIRS

SECOND PART OF THE QUICK FORM

PAIR	YY	%	NY YN	%	NN	%	EP %	DP %	MP %	K	K MAX	K RATIO
1	10	4	12	5	220	91	45	50	95	.60	.93	.64
2	8	3	14	5	242	92	36	58	95	.51	1.00	.51
3	9	4	12	5	221	91	43	52	95	.58	.72	.80
4	19	5	30	8	325	87	39	53	92	.52	.71	.73
5	28	7	54	14	292	78	34	51	86	.44	.56	.78
6	30	8	25	7	308	85	55	39	93	.67	.80	.83
7	33	8	32	8	342	84	51	41	92	.63	.93	.68
8	10	5	29	15	159	80	26	60	85	.34	.52	.65
9	7	3	15	7	198	90	32	61	93	.45	.67	.67
10	22	6	40	10	323	84	35	54	90	.47	.73	.64
11	19	5	32	8	367	88	37	55	92	.50	.81	.62
12	28	7	47	12	321	81	37	51	88	.48	.77	.62
13	32	8	38	9	337	83	46	45	91	.58	.73	.79
14	22	5	30	7	366	88	42	51	93	.56	.85	.65
15	4	1	6	2	375	97	40	58	98	.56	.71	.80

LEGEND
YY= number of times both raters answer "yes"
NYYN= number of times one rater answers "yes" and the other rater answers "no"
NN= number of times both raters answer "no"
EP= effective percent, occurrence percent
DP= mean percent minus effective percent
K= kappa KMAX= kappa max

TABLE 22 TEST-RETEST RELIABILITY OF RATERS

FIRST PART OF THE QUICK FORM

RATER	YY	%	NY YN	%	NN	%	EP %	DP %	MP %	K	K MAX	K RATIO
1	2	5	8	20	30	75	20	60	80	.25	.44	.57
2	11	27	9	22	20	50	55	23	78	.53	.74	.72
3	1	2	1	2	38	95	50	48	98	.66	.66	1.00
4	1	2	5	13	34	85	17	71	88	.25	.25	1.00
5	4	10	6	15	30	75	40	45	85	.48	.83	.58
6	3	7	3	7	34	85	50	43	93	.63	.87	.71
7	11	27	11	27	18	45	50	23	73	.47	.47	1.00
8	7	17	9	22	24	60	44	34	78	.48	.48	1.00
9	1	2	5	13	34	85	17	71	88	.25	.25	1.00
10	5	13	12	30	23	57	29	41	70	.32	.32	1.00
11	10	25	12	30	18	45	45	25	70	.43	.43	1.00
12	10	25	9	22	21	52	53	25	78	.52	.73	.71
13	9	11	17	21	54	67	35	44	79	.39	.75	.52
14	7	17	4	10	29	72	64	26	90	.72	.72	1.00
15	2	7	6	20	22	73	25	55	80	.30	.53	.57
16	3	7	5	13	32	80	38	50	88	.49	.49	1.00
17	4	10	3	7	33	82	57	35	93	.69	.69	1.00
18	4	10	4	10	32	80	50	40	90	.62	.62	1.00
19	15	38	1	2	24	60	94	4	98	.95	.95	1.00
20	1	2	11	27	28	70	8	64	73	.11	.11	1.00
21	10	25	13	32	17	42	43	24	68	.35	.65	.54
22	30	38	13	16	37	46	70	14	84	.67	.87	.77
23	8	20	5	13	27	67	62	26	88	.68	.94	.72
24	13	32	7	17	20	50	65	18	83	.64	.95	.67
25	5	13	3	7	32	80	63	30	93	.73	.73	1.00
26	12	30	3	7	25	63	80	13	93	.83	.83	1.00
27	5	13	0	0	35	88	100	0	100	1.00	1.00	1.00
28	6	15	2	5	32	80	75	20	95	.83	1.00	.83

LEGEND
YY= number of times both raters answer "yes"
NYYN= number of times one rater answers "yes" and the other rater answers "no"
NN= number of times both raters answer "no"
EP= effective percent, occurrence percent
DP= mean percent minus effective percent
K= kappa KMAX= kappa max

TABLE 23 TEST-RETEST RELIABILITY OF RATERS

SECOND PART OF THE QUICK FORM

RATER	YY	%	NY YN	%	NN	%	EP %	DP %	MP %	K	K MAX	K RATIO
1	10	23	3	7	31	70	77	16	93	.82	.82	1.00
2	10	23	4	9	30	68	71	19	91	.77	1.00	.77
3	2	5	3	7	39	89	40	53	93	.54	.54	1.00
4	7	16	5	11	32	73	58	30	89	.66	.93	.71
5	5	11	5	11	34	77	50	39	89	.61	.61	1.00
6	7	16	3	7	34	77	70	23	93	.78	.93	.84
7	17	39	4	9	23	52	81	10	91	.82	.91	.90
8	17	39	4	9	23	52	81	10	91	.82	.91	.90
9	3	7	3	7	38	86	50	43	93	.63	.88	.72
10	8	18	3	7	33	75	73	20	93	.80	.93	.86
11	10	23	9	20	25	57	53	27	80	.55	.65	.85
12	12	27	5	11	27	61	71	18	89	.74	.85	.88
13	25	28	10	11	53	60	71	17	89	.75	.90	.83
14	4	9	2	5	38	86	67	29	95	.78	.78	1.00
15	15	34	4	9	25	57	79	12	91	.81	.90	.89
16	5	11	4	9	35	80	56	35	91	.66	.83	.80
17	1	2	2	5	41	93	33	62	95	.48	.48	1.00
18	9	20	4	9	31	70	69	22	91	.76	.88	.86
19	14	32	5	11	25	57	74	15	89	.76	.86	.89
20	2	5	5	11	37	84	29	60	89	.40	.40	1.00
21	5	11	4	9	35	80	56	35	91	.66	.83	.80
22	30	34	16	18	42	48	65	17	82	.63	.82	.77
23	6	14	5	11	33	75	55	34	89	.64	.93	.69
24	10	23	3	7	31	70	77	16	93	.82	.94	.88
25	7	16	3	7	34	77	70	23	93	.78	.78	1.00
26	6	14	2	5	36	82	75	20	95	.83	1.00	.83
27	1	2	3	7	40	91	25	68	93	.37	.79	.46
28	8	18	1	2	35	80	89	9	98	.93	.93	1.00

LEGEND
YY= number of times both raters answer "yes"
NYYN= number of times one rater answers "yes" and the other rater answers "no"
NN= number of times both raters answer "no"
EP= effective percent, occurrence percent
DP= mean percent minus effective percent
K= kappa KMAX= kappa max

DISCUSSION

VALIDITY

The PBAT is an assessment tool whose initial development differed from that of many instruments of a similar type. A common practice is to base a new tool upon either desirable items of previous ones or upon symptoms thought to measure the construct of interest according to a group of psychologists or psychiatrists. The instrument is then used to gather data of benefit to these professionals, largely either in the assessment of a patient or in the evaluation of pharmacological trials. The PBAT differs in that: (1) its items were elicited from nursing staff responsible for the management of problem behaviors and not from preconceptions regarding behaviors to be included; and (2) it is intended to be useful to nurses in addition to psychologists and physicians. The PBAT's unique method of construction gives it more content validity than many current tools possess. According to Nunnally (1967), content validity is dependent on: 1. the representativeness the item collection, and 2. a sensible method of test construction. The PBAT meets both of these requirements.

The face validity of the PBAT is particularly important in that nursing staff are generally not "test wise" when considering psychological measurements, and are unlikely to complete accurately instruments which appear to have no value or little relevance to their work.

RELIABILITY

The two main types of reliability required by an instrument of this type, inter-rater and test-retest reliability, have been

demonstrated. It should be noted that unlike many such devices, this reliability was based on ratings made according to usual experience with residents, not after viewing the same videotaped interviews (see Kane & Kane, 1981). Reliability has been demonstrated for Quick Form items individually, as opposed to an averaged composite reliability as is frequently reported.

Comparison of the reliabilities of various aspects of the PBAT also suggest it to have a meaningful pattern of internal relationships. These relationships are outlined below:

1. A greater degree of reliability is shown for the more objective and observable items than for the more subjective items, i.e., greater inter-rater reliability is demonstrated for the second part of the Quick Form than for the first part.

2. A greater degree of reliability is shown for frequency estimates than for severity estimates, i.e., raters have greater agreement on a characteristic with an external reality than they do on a characteristic that is a psychological construct. To further this point, for frequency they agree equally well with both themselves (retest) and with their partner (inter-rater). For the correlation of severity with frequency, retest reliability (-.58) is higher than inter-rater reliability (-.38).

3. Generally, test-retest reliability is higher than inter-rater reliability. For example, on severity estimates, raters agree with themselves better than with their partner. However, in some instances, the degree of inter-rater reliability is the higher of the two forms.

POSSIBLE IMPROVEMENTS TO THE PBAT

As experience with a particular tool is gained, awareness of its short-comings increases. These short-comings can range in significance from minor annoyances to major deficits. The present section reviews some of the short-comings identified in using the Problem Behavior Assessment Tool, and outlines possible improvements that might be incorporated if another edition of the PBAT were contemplated.

The reliability of some items on the Quick Form of the PBAT is quite low, and the probable incidence of true positive responses is also low. Such items could be deleted, specifically the items regarding Climbing out of Bed and Bed Restraints. Some Special Scale items are either unimportant or unreliable and could be also be eliminated. An example is the item regarding the disposal of removed objects on the Removing Scale. The elimination of such items has also the obvious benefit of shortening the instrument.

The physical presentation of the PBAT could be improved. Some raters noted that after doing many forms, the line separating the Special Scale items and the line to be bisected for dimensions such as severity became too much alike. This deficit could be remedied easily. Also, the verbal labels attached to the frequency line need improvement. An alternative may be to present a line such as:

FREQUENCY: On how many days a month does this behavior occur?

---

1 15 30

A second alternative would be to simply request the rater to fill in the number of times per month as their response.

Consideration of the Special Scale results suggests that the PBAT appears to ask nurses to make finer behavioral distinctions than they are able to do. It is preferable to begin with finer distinctions than possible, as opposed to the opposite alternative of presenting coarser choices than raters can handle. Either the PBAT could be simplified, or the raters could be better trained.

Additions to the PBAT could also be made. These might include items regarding fighting at lap restraints, and ones examining cyclic behavior, as these are difficulties which I am now becoming more aware of. I am sure other clinicians would have more ideas for new scales.

Other changes that might be worthwhile in a next edition of the PBAT could be the provision of written instructions regarding how to respond in a variety of circumstances. One example is the the more emphatic elimination of physical disabilities. Is a patient with an involuntary movement disorder being physically abusive if his uncontrolled movements result in striking staff? Is the patient with a grasp reflex being physically abusive when he grips the staff? What about the resident who perceives his transfer from wheelchair to bed as being precarious and responds by grabbing at the staff in his fear of falling? Or the resident who has difficulty swallowing and disrupts the dining room with his efforts? Another example is the resident whose body language and spoken text differ. If a resident smiles and has a twinkle in her eye as she calls staff something atrocious, is she being verbally abusive? Some nursing staff say yes, some no.

Instructions and scoring guidelines for circumstances such as these could be developed through a series of sessions with disagreeing raters who discuss their responses and what led to them, hopefully resulting in some resolution. Also desirable would be a means of "giving permission" to say bad things about residents since nurses generally have a strong attachment to them.

USES FOR THE PBAT

There are a variety of ways the PBAT could be used.

The Quick Form of the PBAT allows a rapid and reasonably reliable means of estimating the quantity of behavioral management load on a particular nursing unit. A comparison of data from several nursing units might allow staff and management to conduct negotiations concerning staffing levels with more calm.

It is generally true that nurses currently in practice have received little or no training regarding behavioral observation and documentation. It is possible that the PBAT could be included in a course which aims to remedy this situation.

Once problem behaviors are identified and properly described, the vital task of instituting interventions can begin. A means of quantifying the effectiveness of chosen interventions would be very attractive. The PBAT would be appropriate for such a purpose only under special circumstances. One would especially want to ensure that the particular raters involved were experienced enough with the tool to guarantee its use in a reliable manner.

Recently, the Canadian Ministry of Health and Welfare has announced a Seniors Initiative to promote quality of life for senior

citizens (Epp, 1988). Research activities are to be supported, and a number of areas of particular interest have been specified. One is the investigation of systematic interventions for the management of problem behavior. The results of this dissertation, along with other reports in the literature, would suggest that care be taken in identifying problem behavior residents, and in quantifying characteristics by which change is measured.

NURSES AND NURSING

Though the PBAT is basically reliable, it demonstrates instances where nurses disagree about the typical behaviors of their residents. How does such disagreement come to exist? There may be differential reporting to nurses, i.e., an aide knows that particular nurse is not receptive to hearing of behavioral difficulties or may blame the aide's behavior, so she does not report behavioral difficulties. Nurses may differ in their tolerance of behavior problems; and, given the number of residents they supervise, they may forget which resident shows what kind of behaviors.

Nurses claim that residents display very different behaviors on different shifts. How much of their perceived behavioral difference is due to actual resident behavior and how much to different nurses having different perceptions and different tolerance?

How can one hope to do clinical work with behavior problems if nurses caring for the same residents on the same shift do not agree about their behavior? If nurses on one shift do not agree, how can nurses on different shifts cooperate in conducting a consistent behavioral care management plan? How can they effectively cooperate in

the operation of a consistent behavioral care plan? How frequent is this situation in other areas of clinical care?

Consider the following situation. A resident in a LTC facility has a particular problem behavior, according to a nurse responsible for that resident. A management plan in which nursing staff are not to reinforce the problem behavior is begun. But some nursing staff do not concede that the problem behavior is committed by this resident and hence do not comply with the management plan. The resident is then receiving intermittent reinforcement. This reinforcement schedule develops a very high rate of response. Might this situation be actually occurring frequently in care settings?

The difference between the reliability of the two parts of the Quick Form shows that agreement about the more observable behaviors is greater than agreement about the more subjective behaviors. What contribution does a nurse's personal attitude toward a resident make to her care of that resident?

RELATION TO LITERATURE ON PROBLEM BEHAVIORS AND THEIR ASSESSMENT

Some instruments, such as the PAMIE, are completed by raters who are psychologists or psychiatrists, and contain items about the patient's functioning in the activities of daily living and physical abilities. Yet psychologists or psychiatrists are quite unlikely to have direct knowledge of these areas and will usually either guess or ask care staff. This situation is a source of error. The PBAT is completed by direct care staff who have direct knowledge of the residents whose behavior they are rating.

Cumming, Cumming, Titus, Schmelzle, and MacDonald (1982) conducted a study regarding problem behaviors in an attempt to investigate whether segregated units ought to be provided for behaviorally difficult residents. They asked nurses in LTC, ECU and Home Care placements to nominate problem behavior residents during structured interviews. They asked the question "Do any of your residents have any kinds of behavior problems?" The nurse would mention one resident, there would be a series of questions about the resident and the behaviors, then the interviewer would ask if another resident had problem behaviors, etc. This is a time-consuming procedure, especially for nurses who do not have time to spare - a condition which would tend to limit the number of nominations. Interviews were conducted on two occasions, two months apart. Importantly, the same nurses were not always used on the two occasions. Some of the behaviors elicited are what could be considered more attitudinal or subjective than others, e.g., persistent demands for care and attention, (which was said to be one of the more frequent complaints), resistance to rules, and bizarre behavior, which includes poor mealtime behavior.

The Cummings et al. (1982) study has numerous flaws but some of the most important are the use of different raters on the two interview occasions, not systematically inquiring about all residents in each nurse's care, not attending to subjective versus objective aspects of reported problem behaviors and not checking the reliability of their interview method.

The PBAT was conducted in a more precise manner. Inquires were made about each and every resident. The same raters participated on

both occasions. To nurse's credit, my study shows higher reliability than one would guess to be their the estimate from the Cummings et al. (1981) study.

SAFEGUARDS CONCERNING BEHAVIORAL MANAGEMENT OF LTC RESIDENTS

One final issue remains to be addressed. The PBAT is intended to allow an initial clinical picture to be developed, one which could be used in the formulation and implementation of strategies for the management of identified problem behaviors. Such strategies are likely to draw on the concepts and principles developed for behavioral therapies. When behavioral therapies are considered or proposed, one of the more frequently encountered concerns is that such therapies might be misused, or applied in a manner counter to the client's best interest. The implied recognition of the effectiveness of behavioral therapies is gratifying, but the thought to question the ethics and judgement of practitioners is most wise, and most welcome.

There are two aspects to the issue of improperly applied clinical skills. The first concerns the misuse of psychological tests. Any instrument intended for clinical purposes can be used in a manner counter to the best interests of the client, and the PBAT is no exception. As a safeguard against improper use, commercially published tests are now sold only to individuals recognized as having appropriate credentials. Publicly published materials can offer no such safeguard.

A few statements about the PBAT's potential for misuse can be made. One, the PBAT is an instrument which collects data regarding ongoing behavior. The behaviors examined by the PBAT are ones which most people would find problematic, e.g., verbal abuse, physical abuse,

and wandering. While these behaviors inconvenience the facility staff responsible for their management, they also jeopardize the safety of the resident. Two, the PBAT does not contain any specific recommendations or techniques for altering behavior. Were an unscrupulous operator to intend to change behaviors he found unappealing or undesirable, the PBAT is unlikely to be of direct assistance.

The second aspect concerning the improper application of clinical skills is the clinicians own responsibility for proper practice. McDonald (1982) has said:

> Behavior therapy, like any treatment, should not be applied unilaterally. It should be implemented only following consultation with the recipient, or his or her agent, about whether treatment is justified, whether the specific treatment planned is acceptable, and whether the treatment goals are perceived as beneficial to the person receiving the treatment. While behavioral technology may be used to render people easier to manage, as are medical and physical technologies, such actions cannot be construed as therapeutic; they are mo more a part of behavior therapy than forced sterilization is of medicine. (p. 174)

When I consider devising a behavioral management plan, I like to ensure that four parties participate as willing partners. The first is either the resident him or herself, or agent, such as a family member. The second is the facility staff who ultimately must actually conduct the technique. The third is the facility administration which must condone

the implementation of the technique. The fourth is an independent overseer who can observe what is being done and remain free to question the plan. I hope that the communication and cooperation which this partnership requires provides sufficient care for the benefit of clients.

REFERENCES

Allen, M. & Yen, W. (1979). Introduction to Measurement Theory. Belmont, CA: Wadsworth, Inc.

Bartko, J. J., & Carpenter, W. T. (1976). On the methods and theory of reliability. Journal of Nervous and Mental Disease, 163, 307-317.

Bavelas, A. (1967). Echo: a method for improving understanding between groups. General Research Corp. Santa Barbara. unpublished manuscript.

Bernardin, H.J. & Beatty, R. (1984). Performance appraisal: Assessing human behavior at work. Boston: Kent Publishing Co.

Cohen, J. (1960). A coefficient of agreement for nominal scales. Educational and Psychological Measurement, 20, 3-7.

Cronbach, L. J. (1970). Essentials of psychological testing (3rd ed.). New York: Harper & Row.

Cumming, J., Cumming, E., Titus, J., Schmelzle, & MacDonald, J. (1982). Episodic nature of behavioral disturbances among residents of facilities for the aged. Canadian Journal of Public Health, 73, 319-322.

Dellasega, C. (1986). The 5% fallacy. Journal of Gerontological Nursing, 13, 7.

Epp, J. Federal government to support independence for seniors, news release, February 9, 1988.

Feil, N. (1967). Group therapy with disoriented nursing home residents. Social Work with Groups, 5(2),57-65.

Feil, N. (1982). Group therapy in a home for the aged. Gerontologist, 7(3), 192-195.

Flanagan, J. C. (1949). Techniques for developing critical requirements from critical incidents. American Psychologist, 4, 236.

Fleiss, J. L., & Cohen, J. (1973). Equivalence of weighted kappa and intraclass correlation coefficient measures of reliability. Educational and Psychological Measurement, 33, 613-619.

Folsom, J. C. (1968). Reality orientation for the elderly mental patient. Journal of Geriatric Psychiatry, 1, 291-307.

Grove, W. M., Andreasen, N. C., McDonald_Scott, P., Keller, M. B., & Shapiro, R. W. (1981). Reliability studies of psychiatric diagnosis. Archives of General Psychiatry, 38, 408-413.

Gurel, L., Linn, M. W., & Linn, B. S. (1972). Physical and Mental Impairment-of-Function Evaluation in the Aged: The PAMIE Scale. Journal of Gerontology, 27, 83-90.

Hartmann, D. P. (1977). Considerations in the choice of interobserver reliability measures. Journal of Applied Behavior Analysis, 10, 103-116

Honigfeld, G. (1983). Psychopathology rating scales for use by nursing personnel. In T. Crook, S.Ferris, & R. Bartus (Eds.), Assessment in Geriatric Psychopharmacology (pp. 81-96). New Canaan, Conn.: Mark Powley Associates, Inc.

Honigfeld, G., & Klett, C. (1965). Nurses Observation Scale for Inpatient Evaluation. Journal of Clinical Psychology, 21, 65-71.

Jensen, A. R. (1959). The reliability of projective techniques: methodology. Acta Psychologia, 16, 108-136.

Kane, R. A., & Kane, R. L. (1981). Assessing the elderly: a practical guide to measurement. Toronto: Lexington Books.

Kleban, M. H., Brody, E. M., & Lawton, M. P. (1971). Personality traits in the mentally-impaired aged and their relationship to improvements in current functioning. The Gerontologist, 1, 134-140.

Kratchowill, T. R., & Wetzel, R. J. (1977). Observer agreement, credibility and judgment: some considerations in presenting observer agreement data. Journal of Applied Behavior Analysis, 10, 133-139.

Lawlis, G. F., & Lu, E. (1972). Judgement of Counseling Process: Reliability, agreement and error. Psychological Bulletin, 78, 17-20.

Lu, E. (1971). A measure of agreement among subjective judgments. Educational and Psychological Measurement, 31, 75-84.

Mantle, J. (September 1985). Resident profile survey. in house document

MacDonald, M.L. (1982). Behavioral consultation in geriatric settings. The Behavioral Therapist, 6, 172-174.

Meer, B., & Baker, J. A. (1966). The Stockton Geriatric Rating Scale. Journal of Gerontology, 21, 392-403.

Nunnally, J. (1978). Psychometric Theory. Toronto: McGraw-Hill, Inc.

Pablo, R. Y. (1976). Evaluation of the PAMIE in Long Term Care and rehabilitation hospital patient population. Canadian Journal of Public Health, 67, 305-313.

Plutchik, R., Conte, H., Lieberman, M., Bakur, M., Grossman, J., & Lehrman, N. (1970). Reliability and validity of a scale for assessing the functioning of geriatric patients. Journal of the American Geriatrics Society, 18, 491-500.

Plutchik, R., & Conte, H. (1970). Change in social and physical functioning of geriatric patients over a one-year period, Journal of the American Geriatrics Society, 18, 491-500.

Shader, R. I., Harmatz, J. S., & Salzman, C. (1974). A new scale for clinical assessment in geriatric populations: Sandoz Clinical Assessment - Geriatric (SCAG). Journal of the American Geriatrics Society, 22, 107-113.

Smith, J. M., Bright, B., & McCloskey, J. (1977). Factor analytic composition of the Geriatric Rating Scale, Journal of Gerontology, 32, 58-62.

Smith, P. C. & Kendall, L. (1963). Retranslation of expectations: An approach to the construction of unambiguous anchors for rating scales. Journal of Applied Psychology, 47, 149-155

Spitznagel, E. L., & Helzer, J. E. (1985). A proposed solution to the base rate problem in the kappa statistic. Archives of General Psychiatry, 42, 725-728.

Tinsley, H., & Weiss, D. J. (1975). Interrater reliability and agreement of subjective judgments. Journal of Consulting Psychology, 22, 358-376.

Venn, R.D. (1983). The Sandoz Clinical Assessment - Geriatric (SCAG) Scale. Gerontology, 29, 185-198.

Yelton, A. R., Wildman, B. G., & Erickson, M. T. (1977). A probability based formula for calculating interobserver agreement. Journal of Applied Behavior Analysis, 10, 127-131.

Yesavage, J. A., Adey, M., & Werner, P. D. Development of a geriatric behavioral self-assessment scale. Journal of the American Geriatrics Society, 1981, 29, 285-288.

Neuropsychological Rehabilitation conference, University of Victoria, 1986.

APPENDIX ONE

ELIGIBILITY CRITERIA FOR EXTENDED CARE

A. Characteristics of Extended Care

1. An extended care unit or hospital provides for:

   a. around-the-clock supervision by a graduate nurse as well as

   b. supervision by various other professional health workers such as: pharmacist, dietitian, occupational/physiotherapist, and social worker;

   c. regular medical supervision;

   d. simple nursing procedures once a day or more often, such as application of surgical dressing, administration of injectable medications, oxygen therapy (low concentration, low flow) or catheter care;

   e. fulfillment of social needs of the beneficiary;

   f. a home-like environment;

   g. a program to assist each beneficiary to retain or improve his functional ability;

   h. skilled assistance with activities of daily living, such as dressing, washing, grooming and bathing.

2. In Extended Care, the length of time required by the applicant for skilled and professional health staff services may vary widely but will average over 150 minutes. Professional to non-professional ratio of staff approximately 1:4.

B. Criteria

1. The following criteria shall be used to determine the eligibility of an applicant for extended care; The applicant -

   a. will not at the time require the services of an acute, rehabilitation or psychiatric hospital;

   b. will, in order to be mobile, require human assistance and sometimes also the use of mechanical aids such as braces, walkers, grab bars, canes and crutches (but not articles of furniture) in order to:

i. turn and move about in bed;
ii. transfer and walk with safety a distance of at least 15 feet clear space; or
iii. transfer and operate a wheelchair safely, including use of footpedals and brakes.

c. may be mobile without human assistance but will require, for medical reasons, 24 hour-a-day surveillance by professional health care staff;

d. may be mobile without human assistance but will regularly require the performance of one or more simple specific nursing procedures more often than once daily - for example, giving of injectable medication, the change of surgical dressing, the treatment of pressure sores, the delivery of low-concentration, low-flow oxygen, catheter and ostomy care; or tube feeding. (except when there is also a tracheostomy)

2. In addition to the above criteria for eligibility, applicants may demonstrate the following:

a. Communication - The applicant
(1) may have difficulty expressing needs or be unable to express needs;
(2) may be unable to adapt to visual or auditory losses, for example, a blind person who is also confused.

b. Personal Functions - The applicant
(1) may require a varying amount of assistance with dressing, washing, grooming and bathing.

c. Mental Functions - The applicant
(1) may or may not be mildly depressed or agitated;
(2) may or may not have moderately impaired comprehension (ability to understand only simple instruction, short retention span);
(3) may or may not demonstrate varying degrees of difficulty in orientation as to time, place and persons.

d. Medical Problems - In addition to those problems related to eligibility

The applicant

(1) will require monthly or more frequent visits by a physician;
(2) may require a therapeutic diet;
(3) may require brief periods of individual Physiotherapy or Occupational therapy;
(4) may require professional monitoring and judgement on a continuing basis for a psychiatric problem (may not be available in all Extended Care Units)

(5) may be mobile without human assistance but exhibits gross fecal or urinary incontinence.

e. Social Functions - The applicant
(1) will require a home-like environment;
(2) will require programs for social and recreational activities;
(3) may be attending social or educational facilities outside the Extended Care Unit (including vocational training).

APPENDIX TWO ITEM STATISTICS FIRST EVALUATION

SCALE	QU	ITEM	N	RA	RB	RC	RD	MP	K	KMAX	KRATIO
1	2	1	6	4	1	1	0	66	-.20	1.00	-.20
1	2	2	6	3	1	0	2	83	.67	.67	1.00
1	2	3	6	0	2	0	4	66	*	*	*
1	5	1	6	0	1	1	4	66	-.20	1.00	-.20
1	5	2	6	4	1	1	0	66	-.20	1.00	-.20
1	5	3	6	0	0	0	6	100	*	*	*
1	5	4	6	0	0	2	4	66	*	*	*
1	5	5	6	0	0	2	4	66	*	*	*
1	5	6	6	0	1	2	3	50	-.29	.57	-.50
1	5	7	6	1	1	1	3	66	.25	1.00	.25
2	3	1	8	2	1	2	3	63	.25	.75	.33
2	3	2	8	5	1	1	1	75	.33	1.00	.33
2	4	1	8	3	0	1	4	88	.75	.75	1.00
2	4	2	8	4	2	0	2	75	.50	.50	1.00
2	4	3	8	5	1	0	2	88	.71	.71	1.00
2	4	4	8	0	0	1	7	88	*	*	*
2	5	1	8	1	1	4	2	38	-.11	.33	-.33
2	5	2	8	0	3	1	4	50	-.23	.38	-.60
2	5	3	8	3	3	1	1	50	.00	.50	.00
2	5	4	8	0	0	0	8	100	*	*	*
2	6	1	8	1	1	2	4	63	.14	.71	.20
2	6	2	8	0	2	2	4	50	-.33	1.00	-.33
2	6	3	8	2	4	1	1	38	-.11	.33	-.33
2	6	4	8	1	0	3	4	63	.25	.25	1.00
2	6	5	8	0	1	0	7	88	*	*	*
2	6	6	8	0	0	1	7	88	*	*	*
3	2	1	45	16	9	5	15	69	.38	.82	.46
3	2	2	45	5	7	1	32	82	.46	.59	.77
3	2	3	45	8	14	7	16	53	.06	.69	.09
3	2	4	45	16	7	17	5	47	-.08	.55	-.14
3	2	5	45	13	11	7	14	60	.21	.82	.25
3	2	6	45	1	3	2	39	89	.23	.85	.27
3	4	1	45	4	9	4	28	71	.21	.69	.30
3	4	2	45	41	2	2	0	91	-.05	1.00	-.05
3	4	3	45	12	8	4	21	73	.45	.82	.55
3	4	4	45	3	5	2	35	84	.38	.73	.51
3	5	1	45	43	1	1	0	96	-.02	1.00	-.02
3	5	2	45	29	8	4	4	73	.24	.75	.32
3	6	1	45	38	4	2	1	87	.18	.73	.25
3	6	3	45	14	6	9	16	67	.33	.87	.39
3	6	4	45	20	6	5	14	76	.50	.95	.53
3	6	5	45	4	5	2	34	84	.44	.76	.58
3	6	2	45	16	12	3	14	67	.36	.61	.58
3	7	1	45	11	5	14	15	58	.18	.61	.30
3	7	2	45	6	21	5	13	42	-.05	.35	-.14
3	7	3	45	6	5	6	28	76	.36	.94	.38
3	7	4	45	1	1	0	43	98	.66	.66	1.00
3	8	1	45	18	5	7	15	73	.47	.91	.51
3	8	2	45	16	9	4	16	71	.43	.78	.55

* KAPPAS CAN NOT BE CALCULATED IF TWO CELLS ARE 0.

APPENDIX TWO ITEM STATISTICS - CONTINUED

SCALE	QU	ITEM	N	RA	RB	RC	RD	MP	K	KMAX	KRATIO
4	2	1	54	27	10	5	12	77	.40	.80	.50
4	2	2	54	20	14	6	14	63	.27	.71	.38
4	2	3	54	6	7	5	36	78	.36	.89	.40
4	4	1	54	2	5	2	45	87	.30	.70	.43
4	4	2	54	45	2	5	2	87	.30	.70	.43
4	8	1	54	0	0	0	54	100	*	*	*
4	8	2	54	0	1	1	52	96	.02	1.00	-.02
4	8	3	54	36	2	12	4	74	.24	.46	.53
4	8	4	54	3	12	0	39	80	.27	.27	1.00
5	3	1	14	3	3	1	7	71	.39	.70	.56
5	3	2	14	8	3	2	1	64	.05	.81	.07
5	3	3	14	0	0	0	14	100	*	*	*
5	3	4	14	0	2	0	12	86	*	*	*
5	4	1	14	6	4	3	1	50	-.14	.84	-.17
5	4	2	14	4	2	4	4	57	.16	.72	.22
5	4	3	14	0	0	0	14	100	*	*	*
5	5	1	14	11	1	2	0	71	-.11	.63	-.17
5	5	2	14	2	1	1	10	86	.58	1.00	.58
5	5	3	14	0	0	0	14	100	*	*	*
5	5	4	14	0	0	0	14	100	*	*	*
7	2	1	63	13	2	7	41	86	.65	.80	.80
7	2	2	63	12	9	5	37	78	.47	.85	.56
7	2	3	63	20	5	7	30	79	.60	.93	.65
7	2	4	63	24	5	6	28	83	.65	.97	.67
7	2	5	63	4	4	4	51	87	.43	1.00	.43
7	2	6	63	7	5	1	50	90	.65	.76	.85
7	2	7	63	24	13	8	18	67	.33	.84	.39
7	2	8	63	21	5	11	26	75	.49	.81	.61
7	2	9	63	0	2	0	61	97	*	*	*
7	2	10	63	2	3	3	55	90	.35	1.00	.35
7	2	11	63	3	1	2	57	95	.64	.88	.73
7	2	12	63	5	4	2	52	90	.57	.86	.67
7	2	13	63	13	4	6	40	84	.61	.92	.66
8	1	1	19	0	2	1	16	84	-.08	.64	-.12
8	1	2	19	0	1	2	16	84	-.08	.64	-.12
8	1	3	19	0	3	4	12	63	-.22	.83	-.27
8	1	4	19	6	5	5	3	47	-.08	1.00	-.08
8	1	5	19	2	2	3	12	74	.27	.85	.32
8	1	6	19	0	1	0	18	95	*	*	*
8	2	1	19	0	4	4	11	58	-.27	1.00	-.27
8	2	2	19	0	6	3	10	52	-.27	.58	-.46
8	2	3	19	0	2	7	10	52	-.20	.34	-.58
8	2	4	19	2	5	3	9	58	.04	.76	.05

APPENDIX TWO ITEM STATISTICS - CONTINUED

SCALE	QU	ITEM	N	RA	RB	RC	RD	MP	K	KMAX	KRATIO
9	6	1	10	9	0	1	0	90	*	*	*
9	6	2	10	10	0	0	0	100	*	*	*
9	6	3	10	2	0	0	8	100	1.00	1.00	1.00
9	7	1	10	0	2	0	8	80	*	*	*
9	7	2	10	2	0	3	5	70	.40	.40	1.00
9	7	3	10	5	1	0	4	90	.80	.80	1.00
9	8	1	10	2	1	0	7	90	.74	.74	1.00
9	8	2	10	2	0	4	4	60	.29	.29	1.00
9	8	3	10	4	1	2	3	70	.40	.80	.50
9	8	4	10	3	1	0	6	90	.78	.78	1.00
9	9	1	10	9	0	1	0	90	*	*	*
9	9	2	10	0	0	0	10	100	*	*	*
9	9	3	10	1	0	0	9	100	1.00	1.00	1.00
10	3	1	55	49	2	1	3	85	.64	.88	.73
10	3	2	55	48	6	1	0	87	-.03	.26	-.12
10	3	3	55	21	11	7	16	67	.34	.85	.40
10	4	1	55	2	7	3	43	82	.19	.68	.28
10	4	2	55	1	4	3	47	87	.15	.88	.18
10	4	3	55	4	4	1	46	91	.57	.74	.77
10	4	4	55	1	6	1	47	87	.18	.41	.43
10	5	1	55	5	9	4	37	76	.29	.73	.40
10	5	2	55	0	13	9	33	60	-.24	.77	-.31
10	5	3	55	0	11	6	38	69	-.16	.66	-.25
10	5	4	55	2	11	2	40	76	.14	.40	.35
10	6	1	55	2	6	3	44	84	.22	.74	.30
10	6	2	55	8	2	12	33	75	.38	.56	.69
10	6	3	55	1	3	4	47	87	.15	.88	.18
10	6	4	55	0	3	2	50	91	-.05	.79	-.06
10	7	1	55	9	8	9	29	64	.29	.96	.30
10	7	2	55	31	8	8	8	71	.29	1.00	.29
10	8	1	55	6	8	5	36	76	.33	.85	.39
10	8	2	55	9	10	12	24	60	.14	.92	.15
10	8	3	55	14	9	10	22	65	.29	.96	.31

APPENDIX THREE ITEM STATISTICS - SECOND EVALUATION

SCALE	QU	ITEM	N	RA	RB	RC	RD	MP	K	KMAX	KRATIO
1	2	1	18	13	0	1	2	83	.49	.49	1.00
1	2	2	18	7	3	0	8	83	.67	.67	1.00
1	2	3	18	1	2	1	14	83	.31	.77	.40
1	5	1	18	1	2	2	13	78	.20	1.00	.20
1	5	2	18	9	4	0	5	78	.56	.56	1.00
1	5	3	18	1	0	1	16	94	.64	.64	1.00
1	5	4	18	1	1	2	14	83	.31	.77	.40
1	5	5	18	3	1	2	12	83	.56	.85	.65
1	5	6	18	3	2	4	9	67	.26	.75	.35
1	5	7	18	1	2	1	14	83	.31	.77	.40
2	3	1	20	7	2	2	9	80	.60	1.00	.60
2	3	2	20	6	1	3	10	80	.59	.79	.74
2	4	1	20	14	0	0	6	100	1.00	1.00	1.00
2	4	2	20	8	4	1	7	75	.51	.71	.72
2	4	3	20	7	0	3	10	85	.70	.70	1.00
2	4	4	20	14	3	2	1	75	.14	.83	.17
2	5	1	20	3	6	8	3	30	-.39	.80	-.48
2	5	2	20	12	6	2	0	60	-.18	.41	-.43
2	5	3	20	6	2	8	4	50	.07	.44	.17
2	5	4	20	20	0	0	0	100	*	*	*
2	6	1	20	9	1	2	8	85	.70	.90	.78
2	6	2	20	13	3	1	3	80	.47	.74	.64
2	6	3	20	9	2	2	7	80	.60	1.00	.60
2	6	4	20	13	1	2	4	85	.63	.87	.71
2	6	5	20	20	0	0	0	100	*	*	*
2	6	6	20	19	0	1	0	95	*	*	*
3	2	1	45	21	6	2	16	82	.64	.82	.78
3	2	2	45	40	0	1	4	98	.88	.88	1.00
3	2	3	45	24	6	2	13	82	.62	.81	.77
3	2	4	45	13	6	1	25	84	.67	.76	.88
3	2	5	45	25	2	10	8	73	.40	.60	.67
3	2	6	45	37	0	5	3	89	.50	.50	1.00
3	4	1	45	27	5	4	9	80	.52	.95	.55
3	4	2	45	1	3	3	38	87	.18	1.00	.18
3	4	3	45	20	1	7	17	82	.65	.74	.88
3	4	4	45	43	0	1	1	98	.66	.66	1.00
3	5	1	45	1	5	1	38	87	.20	.46	.42
3	5	2	45	6	2	3	34	89	.64	.93	.69
3	6	1	45	3	5	7	30	73	.17	.86	.20
3	6	2	45	23	3	8	11	76	.48	.76	.63
3	6	3	45	22	6	9	8	67	.27	.85	.31
3	6	4	45	18	5	6	16	76	.51	.96	.53
3	6	5	45	27	5	8	5	71	.25	.83	.30
3	7	1	45	20	5	10	10	67	.31	.77	.40
3	7	2	45	14	7	6	18	71	.42	.96	.44
3	7	3	45	28	5	2	10	84	.63	.84	.75
3	7	4	45	45	0	0	0	100	*	*	*
3	8	1	45	15	8	6	16	69	.38	.91	.42
3	8	2	45	15	6	8	16	69	.38	.91	.42

APPENDIX THREE ITEM STATISTICS - TIME TWO - CONTINUED

SCALE	QU	ITEM	N	RA	RB	RC	RD	MP	K	KMAX	KRATIO
4	2	1	29	6	5	8	10	55	.10	.79	.12
4	2	2	29	6	8	3	12	62	.23	.65	.36
4	2	3	29	21	1	3	4	86	.58	.79	.74
4	4	1	29	27	0	0	2	100	1.00	1.00	1.00
4	4	2	29	2	0	0	27	100	1.00	1.00	1.00
4	8	1	29	28	1	0	0	97	*	*	*
4	8	2	29	27	0	1	1	97	.65	.65	1.00
4	8	3	29	2	3	0	24	90	.52	.52	1.00
4	8	4	29	25	1	3	0	86	-.05	.47	-.12
5	3	1	23	9	1	1	12	91	.82	1.00	.82
5	3	2	23	6	2	4	11	74	.46	.82	.56
5	3	3	23	19	0	0	4	100	1.00	1.00	1.00
5	3	4	23	16	2	3	2	78	.31	.86	.36
5	4	1	23	2	3	1	17	83	.40	.70	.57
5	4	2	23	14	1	2	6	87	.70	.90	.78
5	4	3	23	23	0	0	0	100	*	*	*
5	5	1	23	1	1	0	21	96	.65	.65	1.00
5	5	2	23	16	0	3	4	87	.65	.65	1.00
5	5	3	23	23	0	0	0	100	*	*	*
5	5	4	23	23	0	0	0	100	*	*	*
7	2	1	32	16	3	1	12	88	.75	.87	.86
7	2	2	32	20	1	5	6	81	.55	.70	.78
7	2	3	32	11	4	6	11	69	.38	.88	.43
7	2	4	32	12	5	4	11	72	.44	.94	.47
7	2	5	32	27	1	2	2	91	.52	.84	.62
7	2	6	32	27	0	3	2	91	.53	.53	1.00
7	2	7	32	18	3	4	7	78	.50	.93	.54
7	2	8	32	19	2	3	8	84	.65	.93	.70
7	2	9	32	31	0	1	0	97	*	*	*
7	2	0	32	30	0	1	1	97	.65	.65	1.00
7	2	1	32	23	2	3	4	84	.52	.90	.57
7	2	2	32	21	2	1	8	91	.78	.93	.84
7	2	3	32	13	4	4	11	75	.50	1.00	.50
8	1	1	20	15	4	1	0	75	-.09	.35	-.25
8	1	2	20	0	1	4	15	75	-.09	.35	-.25
8	1	3	20	12	4	2	2	70	.21	.74	.29
8	1	4	20	5	3	6	6	55	.12	.71	.17
8	1	5	20	14	2	2	2	80	.38	1.00	.38
8	1	6	20	19	1	0	0	95	*	*	*
8	2	1	20	12	1	1	6	90	.78	1.00	.78
8	2	2	20	15	1	1	3	90	.69	1.00	.69
8	2	3	20	16	1	3	0	80	-.08	.46	-.18
8	2	4	20	11	2	0	7	90	.79	.79	1.00

APPENDIX THREE ITEM STATISTICS - TIME TWO - CONTINUED

SCALE	QU	ITEM	N	RA	RB	RC	RD	MP	K	KMAX	KRATIO
9	6	1	18	1	0	0	17	100	1.00	1.00	1.00
9	6	2	18	3	0	2	13	89	.68	.68	1.00
9	6	3	18	17	0	0	1	100	1.00	1.00	1.00
9	7	1	18	10	1	3	4	78	.51	.75	.67
9	7	2	18	11	4	0	3	78	.48	.48	1.00
9	7	3	18	10	0	2	6	89	.77	.77	1.00
9	8	1	18	9	3	3	3	67	.25	1.00	.25
9	8	2	18	5	5	3	5	56	.12	.78	.16
9	8	3	18	8	4	2	4	67	.31	.77	.40
9	8	4	18	12	0	3	3	83	.57	.57	1.00
9	9	1	18	1	0	0	17	100	1.00	1.00	1.00
9	9	2	18	18	0	0	0	100	*	*	*
9	9	3	18	17	0	0	1	100	1.00	1.00	1.00
0	3	1	46	2	0	3	41	93	.54	.54	1.00
0	3	2	46	3	0	2	41	96	.73	.73	1.00
0	3	3	46	16	3	2	25	89	.77	.95	.81
0	4	1	46	35	6	3	2	80	.20	.73	.27
0	4	2	46	34	7	4	1	76	.02	.73	.03
0	4	3	46	40	4	1	1	89	.24	.54	.44
0	4	4	46	34	7	5	0	74	-.15	.81	-.18
0	5	1	46	25	5	7	9	74	.41	.90	.45
0	5	2	46	28	5	12	1	63	-.09	.55	-.16
0	5	3	46	37	3	4	2	85	.28	.90	.31
0	5	4	46	32	5	3	6	83	.49	.87	.56
0	6	1	46	35	6	3	2	80	.20	.73	.27
0	6	2	46	31	5	4	6	80	.45	.94	.47
0	6	3	46	36	6	2	2	83	.25	.62	.39
0	6	4	46	39	6	1	0	85	-.04	.26	-.15
0	7	1	46	21	7	4	14	76	.51	.87	.59
0	7	2	46	14	3	6	23	80	.59	.86	.69
0	8	1	46	30	3	4	9	85	.62	.95	.65
0	8	2	46	22	9	3	12	74	.46	.73	.63
0	8	3	46	25	3	7	11	78	.52	.81	.65

* KAPPAS CAN NOT BE CALCULATED IF TWO CELLS ARE 0.

APPENDIX FOUR PBAI SECOND VERSION

PROBLEM BEHAVIOR ASSESSMENT INSTRUMENT
SECOND VERSION

developed by:
KATHY MONTGOMERY

Department of Psychology
University of Victoria
P O Box 1700
Victoria, B C V8W 2Y2

PROBLEM BEHAVIOR ASSESSMENT INSTRUMENT

INTRODUCTION

The Problem Behavior Assessment Instrument (PBAI) is a rating scale designed identify and quantify problem behaviors exhibited by residents of Long Term Care Units. The problem behaviors included on the scale are those identified by direct-care nurses as being ones which hampered the provision of quality care.

The PBAI has two main components:

1. The first main component is entitled the Survey Form. The Survey Form contains two types of questions. The first type concerns irritating questions, and is simply answered "yes" or "no". The second type concerns more problematic behaviors and is also answered "yes" or "no", but the "yes" answer is accompanied by an instruction to complete a special scale (located in the second main component) regarding that problem behavior.

2. The second main component is a collection of separate Special Scales, each regarding a specific problem behavior. These Special Scales correspond to problem behaviors from the Survey Form, and require specifying characteristics, such as severity, frequency, etc. Only scales for problem behaviors actually observed in a given resident are completed for that resident; the other scales are omitted.

By having these two main components, the PBAI is as short as possible while still being specific. For example, a resident who displays only the more minor irritating behaviors and no problem behaviors will only require the Survey Form; whereas the resident who displays three problem behaviors will require the Survey Form plus the corresponding three special scales.

The Survey Form should be answered completely before moving onto the Special Scales.

PROBLEM BEHAVIOR ASSESSMENT INSTRUMENT

RESIDENT'S NAME _____ DATE _____
(day/month/year)

RATER'S NAME _____

FACILITY NAME _____

Questions in this section are somewhat general and broad, as for example, "Is the resident verbally abusive?" These questions should be answered according to the characteristic typical behavior of the resident as you have come to know him or her. Answer according to the typical behavior, not according to an unusual incident.

Consider only the resident's current behavior in the last four weeks.

For example, if the question is "Is the resident verbally abusive?" and the resident you are rating becomes nasty, hostile and swears regularly, you would answer "yes". Conversely, if you are rating a resident who required and unpleasant treatment and responded by being verbally abusive for the only time you are ware of, you would answer "no".

All questions on the Survey Form should be answered before filling out the Special Scales.

PROBLEM BEHAVIOR ASSESSMENT INSTRUMENT

SURVEY FORM: PART ONE
Circle an answer for each of the questions below.

---

1. Is the resident considered demanding, "spoiled" or manipulative by many staff members? YES NO

---

2. Does the resident insist on an excessive amount of assistance considering his or her mental and physical abilities?

YES NO

Is the resident uncooperative or resistant to:

3. morning care, H. S. care YES NO
4. toileting YES NO
5. nutrition YES NO
6. transferring YES NO
7. medication administration YES NO
8. treatments YES NO

---

9. Is the resident generally irritable, grouchy, unpleasant?

YES NO

---

10. Does the resident make complaints that the staff generally believe to be unjustified? YES NO

---

11. Are there problems concerning the resident's mealtime behavior?

YES NO

PROBLEM BEHAVIOR ASSESSMENT INSTRUMENT

SURVEY FORM: PART TWO
Mark an answer for each of the questions below.

---

1. Does the resident rummage through other's things? (Rummage means going through other's things - looking through drawers, cupboards, etc.)

   YES (answer questions on scale 1)
   NO

---

2. Does the resident remove things inappropriately? (Remove things means taking things away from where they are kept when the resident does not own them or need to use them.)

   YES (answer questions on scale 2)
   NO

---

3. Is the resident verbally abusive?

   YES (answer questions on scale 3)
   NO

---

4. Is the resident physically abusive to staff?

   YES (answer questions on scale 4)
   NO

---

5. Does the resident physically abuse other residents?

   YES (answer questions on scale 5)
   NO

---

6. Does the resident wander?
   Wandering is defined as attempting to go to, or being in, a place where the individual resident is not allowed. Wandering does not include aimless moving about within proper areas.

   YES (answer questions on scale 6)
   NO

---

7. Does the resident create disruptive unnecessary noise? (screaming, moaning, etc.)

   YES (answer questions on scale 7)
   NO

Circle the scales to be completed:
1 2 3 4 5 6 7
Now answer the questions on the scales circled above.

INSTRUCTIONS FOR THE SPECIAL SCALES

Complete only the special scales which correspond to questions answered "yes" on the Survey Form.

The Special Scales contain questions written in two formats. One format presents a question which is answered by checking off one or more items listed below it. The second format presents a question followed by a line which has words printed along it. These questions are answered by putting a vertical slash through the line. DO NOT CIRCLE THE WORD. Scores for these questions are determined by measuring the line from its far left to the slash mark.

The first question on each scale asks how serious a problem a given behavior is. Answers range from "not a problem" to "wishing the resident could be transferred". Answers in the "severe" or "wishing the resident could be transferred" areas reflect the strain placed on the unit as a whole by a resident's difficult behavior. Answers in these areas do not label a resident as "bad"; they give an honest reflection of the difficulty encountered in trying to manage some problem behaviors displayed by residents despite everyone's best efforts.

RUMMAGING - SCALE ONE

Rummaging means "going through" other's things, e.g., looking through drawers, etc. without necessarily taking things away.

---

1. How serious a problem is the resident's rummaging?

not a problem	mild	moderate	severe	wish the resident would be transferred

---

2. On how many days a month does the resident rummage?

1	15	30

---

3. Where does the resident like to rummage?

_____ In other resident's things
_____ In innocuous things belonging to the unit
_____ In potentially dangerous areas, e.g., med cart, cleaning supplies

---

4. Does the resident stop rummaging readily at staff's request or intervention?

stops readily	may stop	does not stop without staff's insistent effort

---

COMMENTS

REMOVING - SCALE TWO

Remove - means that the item is taken away from where it is kept, does not belong to the resident, and is intended not for his use.

---

1. How serious a problem is the resident's removing?

not a problem	mild	moderate	severe	wish the resident would be transferred

---

2. On how many days a month does the resident remove things?

1 ——— 15 ——— 30

---

3. What does the resident remove? Mark any that apply.

_____ Edible items, e.g., food, drink, candy
_____ Other resident's personal possessions, e.g., clothing, cosmetics, grooming aids, teeth
_____ Items belonging to the unit
_____ Other (please specify)

COMMENTS

VERBAL ABUSE - SCALE THREE

---

1. How serious a problem is the resident's verbal abuse?

not a problem	mild	moderate	severe	wish the resident would be transferred

---

2. On how many days a month is the resident verbally abusive?

1 ——— 15 ——— 30

---

3. What is the nature of the verbal abuse?

_____ Swears and curses, e.g., hell, damn, etc.
_____ Obscene and vulgar langauge, e.g., words with a sexual meaning
_____ Makes verbal threats and accusations
_____ Name calling
_____ Harassment, belligerency, tongue lashing
_____ Other (please specify)

---

4. Whom does the resident direct verbal abuse towards? Mark any that apply.

_____ Abuse is not directed toward anyone specifically - resident just lets out verbal abuse, could even be alone
_____ Directed toward staff
_____ Directed toward other residents

---

5. Where is the resident when being verbally abusive? Mark nay that apply.

_____ Private areas, e.g., bedroom, bathroom
_____ Common areas, e.g., dining room, hallway

---

6. When is the resident verbally abusive?

_____ When receiving daily physical assistance, e.g., dressing, toileting
_____ When receiving a treatment
_____ When in pain or in physical discomfort
_____ During social interactions

COMMENTS

PHYSICAL ABUSE TO STAFF - SCALE FOUR

---

1. How serious a problem is the resident's physical abuse of the staff?

not a problem	mild	moderate	severe	wish the resident would be transferred

---

2. On how many days a month does the resident abuse staff?

1	15	30

---

3. In what way(s) is the resident physically abusive to staff?

_____ Minor - e.g., spits, pushes, scratches, pulls hair
_____ Moderate - e.g., bites, kicks, slaps, twists arm
_____ Severe - e.g., tries to choke, punches, uses weapon

---

4. Has the resident caused a WCB injury to a staff member within the last year?

_____ yes
_____ not to my knowledge

---

5. What is the resident's size?

Height:

short	average	tall

Weight:

light	average	heavy

---

6. What are the resident's physical abilities?

Strength:

weak	average	strong

Mobility:

_____ Walks without assistance
_____ Uses walker without assistance
_____ Uses wheelchair independently
_____ Unable to move about without assistance

COMMENTS

PHYSICAL ABUSE DIRECTED TOWARD OTHER RESIDENTS - SCALE FIVE

1. How serious a problem is the resident's physical abuse of other residents?

not a problem	mild	moderate	severe	wish the resident would be transferred

2. On how many days a month does the resident physically abuse other residents?

1	15	30

3. What is the consequence of the abuse for the recipient?

_____ Emotionally upset, without physical consequence
_____ Bruise, scratch or similar result
_____ Medical care required
_____ Sent to acute hospital

COMMENTS

WANDERING - SCALE SIX

Wandering is defined as attempting to go to, going to, or being in a place where the individual resident is not allowed.
Wandering does not include aimless moving about within proper areas.

---

1. How serious a problem is the resident's wandering?

not a problem	mild	moderate	severe	wish the resident would be transferred

---

2. On how many days a month does the resident wander:

Into other resident's rooms?

0 ——— 1 ——— 15 ——— 30

In other areas on the unit?

0 ——— 1 ——— 15 ——— 30

Off the unit (but still in the building)?

0 ——— 1 ——— 15 ——— 30

Outside the building?

0 ——— 1 ——— 15 ——— 30

---

3. On what shifts does the resident wander?

_____ Days _____ Eves _____ Nights

---

4. How difficult is it to prevent or limit this resident's wandering?

_____ Less difficult than most wanderers
_____ Similar to most wanderers
_____ More difficult than most wanderers

(continued on the next page)

(WANDERING - SCALE SIX continued)

---

5. Once the resident has wandered, what is his/her response to staff attempts to return him/her to a proper area?

_____ Cooperates
_____ Makes verbal objections only
_____ Physically resistive to being returned
_____ Physically aggressive to staff

---

6. Is this resident wandering:

_____ in a wheelchair?
_____ with a walker?
_____ by walking?

COMMENTS

NOISE - SCALE SEVEN

---

1. How serious a problem is the unnecessary noise created by the resident?

not a problem	mild	moderate	severe	wish the resident would be transferred

---

2. On how many days a month does the resident create disruptive unnecessary noise?

1	15	30

---

3. On what shifts does this occur?

_____ Days _____ Eves _____ Nights

---

4. What is the nature of the noise, and how loud is it?

	Audible		
	in resident's room only	in restricted part of unit	throughout unit
repetitive words (help me, help me)	_____	_____	_____
yelling, screaming, moaning	_____	_____	_____
crying	_____	_____	_____
other	_____	_____	_____

COMMENTS

CURRICULUM VITA

1989

KATHLEEN M. MONTGOMERY

Date of birth: April 30, 1953 Marital status: Married
Citizenship: USA, landed immigrant in Canada

EDUCATION AND DEGREES

M. A. 1982 University of Victoria
Dept. of Psychology
Thesis title: "A Normative Study of Neuropsychological Test Performance of an Normal Elderly Sample"

B. A. 1975 University of Massachusetts
Dept. of Psychology

ACADEMIC AWARDS

Medical Research Council of Canada Studentship, 1983/84, 1984/85, 1985/86, 1986/87
University of Victoria Fellowship, 1980/81, 1981/82, 1982/83
Ruth A. Gardner Award, 1985

PROFESSIONAL EXPERIENCE

Psychologist, 1988 -
Victoria Mental Health Center, Wendell Coady, Director

Psychology Technician 1976-1980
Boston Veterans Administration Medical Center, Psychology Department. Supervisor - Nelson Butters, Ph.D.

Research Assistant 1974-1976
Boston University School of Medicine, Dept. of Neurology.
Supervisor - Nelson Butters, Ph.D.

PRACTICUM

Counselling Center, University of Victoria, 1983/84

INTERNSHIP

January 1, 1987 - May 30, 1987:
Juan de Fuca Hospitals, Victoria, B. C.
Valleyview Hospital, Port Coquitalam, B. C.

CONFERENCE PRESENTATIONS

Montgomery, K. Case study: Reduction of screaming by a Long-Term Care resident. Canadian Psychological Association, Halifax, June, 1989

Montgomery, K. A problem behavior assessment tool for Long-Term Care. International Neuropsychological Society Meeting, Vancouver, February, 1989.

Montgomery, K. Development of a problem behavior assessment tool. Canadian Psychological Association Annual Meeting, Toronto, June, 1986.

Montgomery, K. and Costa, L. Concurrent validity of the Mattis Dementia Rating Scale. Paper presented at the International Neuropsychological Society Meeting, Lisbon, Portugal, June 1983.

Montgomery, K. and Costa, L. Neuropsychological test performance of a normal elderly sample. Paper presented at the International Neuropsychological Society Meeting, Mexico City, February 1983.

Jones, B. P., Butters, N., Moskowitz, H. R., and Montgomery, K. Psychophysical assessment of the sensory capacities of Korsakoff, right hemisphere, and non-amnesic alcohol patients. Paper presented at the International Neuropsychological Society Meeting, Santa Fe, February 1977.

PUBLICATIONS

Montgomery, K. Development of a problem behavior assessment tool. Canadian Gerontological Nurse, 2(2):3, 1986.

Sax, D.S., O'Donnell, B., Butters, N., Menzer, L., Montgomery, K., and Kayne, H. L. Computed tomographic, neurologic and neuropsychological correlates of Huntington's Disease. International Journal of Neuroscience, 18, 21-36, 1983.

MacVane, J., Butters, N., Montgomery, K., and Farber, J. Cognitive functioning in male social drinkers. Journal for Studies on Alcohol, 43(1), 81-95, 1982.

Ryan, C., Butters, N., Montgomery, K., Adinolfi, A., and Didario, B. Memory deficits in chronic alcoholics: continuities between the intact alcoholic and the alcoholic Korsakoff patient. In H. Begleiter and B. Kissin (Eds.), Alcohol Intoxication and Withdrawal. Plenum Press, New York, 1982.

Butters, N., Sax, D., Montgomery, K., and Tarlow, S. Comparisons of the neuropsychological deficits associated with early and advanced Huntington's Disease. Archives of Neurology, 35, 585-589, 1978.

Meudell, P., Butters, N., and Montgomery, K. Role of rehearsal in the short-term memory performance of patients with Korsakoff's and Huntington's disease. Neuropsychologia, 16, 507-510, 1978.

Jones, B. P., Butters, N., Moskowitz, H. R., and Montgomery, K. Olfactory and gustatory capacities of alcoholic Korsakoff patients. Neuropsychologia, 16, 323-337, 1978.

Butters, N., Cermak, L. S., Montgomery, K., and Adinolfi, A. Some comparisons of the memory and visuoperceptive deficits of chronic alcoholics and patients with Korsakoff's disease. Alcoholism: Clinical and Experimental Research, 1, 73-80, 1977.

PARTIAL COPYRIGHT LICENSE

I hereby grant the right to lend my dissertation to the users of the University of Victoria Library, and to make single copies only for such users or in response to a request from the Library of any other university, or similar institution, on its behalf or for one of its users. I further agree that permission for extensive copying of this dissertation for scholarly purposes may be granted by me or a member of the University designated by me. It is understood that copying or publication of this dissertation for financial gain shall not be allowed without my written permission.

THE DEVELOPMENT OF A PROBLEM BEHAVIOR ASSESSMENT TOOL

Author

[redacted]

KATHY MONTGOMERY

March 29 1989

DATE