DEVELOPMENT AND ASSESSMENT
OF AN ALTERNATIVE TO THE NARRATIVE OBSERVATION

A Thesis

Presented to

the Faculty of the Department of Psychology

University of Houston

---

in Partial Fulfillment

of the Requirements for the Degree

Master of Arts

---

by

Lee R. Crowley

May, 1976

DEVELOPMENT AND ASSESSMENT
OF AN ALTERNATIVE TO THE NARRATIVE OBSERVATION

An Abstract of a Thesis

Presented to

the Faculty of the Department of Psychology

University of Houston

---

in Partial Fulfillment

of the Requirements for the Degree

Master of Arts

---

by

Lee R. Crowley

May, 1976

ABSTRACT

For seven years, the narrative observation system has been the principal tool of the Behavioral Ecology Research Team in its study of health care at the Texas Institute for Rehabilitation and Research. When the program of research was begun, the narrative observation approach to data collection and analysis (separation of observation from coding and data reduction) was chosen because nothing was known in an explicit and quantitative sense about the behavioral trajectory of spinal cord injured persons. Although costly in terms of necessary equipment and person hours, this approach proved highly successful. The treatment staff has asked to receive the most up-to-date behavioral measures possible for their patients, so that they can use the data to plan patient programs.

The purpose of this study was to develop an alternative system for data collection and management that is more efficient than the narrative system, yet which yields the same derived measures of interest. A pencil-and-paper shortform (which combines and eliminates various steps) was devised for observers. Two sets of ten paired observations were completed. The first set paired ten narrative observations with ten shortform observations in an attempt to assess the convergent validity of the two systems. The second set paired shortform observations in an attempt to assess the reliability of the new system.

At several levels of detail, the shortform system displayed a great deal of congruence with the narrative system and the shortform also proved to be very reliable. The shortform also afforded considerable savings in person hours and other resources. Methodological implications, recommendations for other applications, and guidelines for further research are discussed.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

LIST OF TABLES

LIST OF TABLES (Continued)

LIST OF FIGURES

CHAPTER I

INTRODUCTION

The present study must be understood within its larger context--the program of research on patient behavior and rehabilitation conducted by the Behavioral Ecology Research Team (Willems, in press). This ongoing program involves a number of research projects that are all coordinated around a central set of themes. In 1968, the Texas Institute for Rehabilitation and Research (TIRR) asked Willems and his associates to devise techniques and measures which could be used in the evaluation of the hospital's program of comprehensive rehabilitation for spinal cord injured adults. Until that time, very little was known in an explicit and quantitative sense about the behavioral trajectory of spinal cord injured persons.

When a person incurs a spinal cord injury that renders him quadriplegic or paraplegic, many different problems arise. Providing him with effective rehabilitation services places complicated demands on a system of health care. Immediately after the injury, the person is often dependent upon intervention and mediating care in order to survive. Soon a host of physiological, neural, and anatomical complications arise; e.g., urinary, muscular, circulatory, and hormonal problems that must be remedied. Next, a whole family of behavioral issues arise. With a high cervical-level spinal cord injury (quadriplegia), nearly all of the person's repertoire of everyday behaviors and performances has been eliminated--so much so that verbal interchange with the environment is nearly all that remains in the repertoire of adaptive skills and behaviors.

The central problem in rehabilitation for spinal cord patients is to create conditions under which as much as possible of that repertoire will be replaced, restored, or supplanted. In fact, it can be argued that the focal problem of rehabilitation for spinal cord injury is to restore or substitute as much as possible of the injured person's lost or altered repertoire, to train him to use alternative ways of relating to his environment, and to help him alter his environment in appropriate ways. The goals of rehabilitation are usually stated in behavioral terms; e.g., "to help the person become maximally independent." "Independence" is a behavioral issue because it points to performances which the person can carry out in his usual environment with a minimum of intervention and support from others. In these terms, rehabilitation means intensive and goal-oriented addition, rearrangement, and substitution in the patient's repertoire of behavior and behavior-environment relations.

In addition to being heavily oriented toward functional performance, the perspective and focus of rehabilitation must also be longitudinal. The evolution of the disabled person's repertoire of performance often occurs gradually over time, and the transition to resettlement in the world outside the hospital is also a gradual, lengthy process. Much of the person's time and effort is devoted to carrying out functions and procedures he has been taught, and his well-being depends upon his ability to apply and practice what he has been taught. Thus, the rehabilitation process is a very long one, extending for months or years. Given the importance of what the disabled person does and uses in his everyday environment over a long period of time, it follows that assessment of his status at any given time and his progress or regression over time should focus on his actual performance.

Ironically, even though behavior and performance are so focal to rehabilitation that they provide the philosophical core of rehabilitative care, very little is known in a formal and explicit sense about the <u>behavioral</u> aftermath of spinal cord injury. A search through the rehabilitation literature in 1968 yielded a number of anatomical, physiological, and neurological measures that could be used to assess the spinal cord injured person's physical restoration. Yet, no techniques or instruments were found to assess the amount, adequacy, or integrity of his behavioral restoration in any complete, ongoing way. These deficiencies are largely still present; even though behavioral readjustment is the goal, rehabilitation professionals have not developed dependable, longitudinal, behavioral indicators of patient progress. Just as importantly, these professionals seldom conduct behavioral evaluations of their programs.

The program undertaken by Willems and the Behavioral Ecology Research Team is an attempt to fill these voids in the rehabilitation field by conducting an extensive longitudinal analysis of patient behavior. The precept of the study is to document patient behavior on a relatively continuous basis from admission to TIRR until discharge (an average of three months) and then beyond discharge for 12 months. The perspective and strategy of this research program are derived in part from an ecological view of human behavior and adaptation (Barker, 1965, 1968; Willems, 1973, in press). The focus is on one of the critical domains of health status evaluation--functional performance of target patients--and one of the purposes is to develop and test a new method for assessing health status and change (Willems, in press). Functional performance refers to what patients actually do over extended periods of time. Performance assessment is concerned with measuring directly the when, where, and how of what patients do, quantifying the number and diversity of specific behaviors a patient engages in, and the rates of change in these measures over time. For health care professionals in rehabilitation, functional performance and adaptation of patients are central issues, and they represent the critical pay-offs for the health care system.

The longitudinal analysis of patient behavior presently being conducted by the Behavioral Ecology Research Team (Willems, 1975, in press) is much more heavily oriented toward pay-offs in impact, intervention, and evaluation that its antecedent projects (LeCompte & Willems, 1970; Vineberg & Willems, 1971; Willems, 1972a, 1972b; Willems, Crowley, & Dreher, 1974; Willems & Vineberg, 1969a, 1969b, 1970). There were no precedents for the behavioral study of rehabilitation prior to this series of studies, so an extensive developmental phase was required in which procedures, categories of data, methods of analysis, and issues of reliability could be designed and tested. When the program of research was first begun, a direct narrative observation approach to data collection was chosen because so little was known about the behavioral trajectory of spinal cord injured persons. The direct observation of 12 patients for one day each in the summer of 1968, a similar study of 15 patients in 1971, the antecedents and accompaniments of these efforts, volunteer projects, and a number of graduate student theses and dissertations have all contributed to the background of the current longitudinal study.

Partly on conceptual grounds and partly on the basis of empirical demonstration, it has become an accepted canon of behavioral science that direct observation of persons in their everyday routines (rather than in experimental laboratories and testing situations) is the most direct path toward documenting human performance and understanding the process of human behavioral adaptation to the environment (Barker, 1964; Johnson & Bolstad, 1973; Weick, 1968; Willems, 1973; Willems & Campbell, 1975; Willems & Raush, 1969; Wright, 1960, 1967). The direct observation approach to data collection was chosen for the study of spinal cord injured patients at TIRR because it provides the most straightforward and extensive picture of what patients do. Altmann (1973) calls this approach focal-subject sampling, and Wright (1960) calls it specimen observation. In this approach, a single target patient is observed for a period of time and a record describing his behavior sequence and situation is preserved for coding. Weick (1968) suggests that observers who devote their efforts to collecting such extensive accounts of their target's behavior are typically guided by the empirical approach, and states that

> ...while this approach increases the investigator's chances that he is looking at something stable, there is a greater risk that he will not know what that something is. This risk is partly offset by the fact that most observational records are rich in detail, and this often helps the investigator induce an explanation from his findings (p. 402).

The methodology used in the study of spinal cord injured persons at TIRR differs in important ways from traditional studies of handicapped persons. Traditional psychological studies focus on structures of personality, thoughts, feelings, moods, motives, perceptions, and a host of other traits and dispositions. The anchoring point for the work of Willems and the Behavioral Ecology Research Team is the molar performances of persons vis-a-vis the environment and their responses to it; e.g., where persons go, what they do with their time, how they interact with each other. Willems (in press) argues that

> ...behavioral performance and response are the major means by which person-environment adaptations are actually mediated. The ecologist does not argue against the reality of subjective phenomena. Rather, his position is a matter of emphasis in seeking to understand human functioning in relation to the environment. Nor does the relative emphasis on behavior imply a simplistic advocacy of behaviorism. To the ecologist, overt behavior simply is more important than many other psychological phenomena. To the behavioral ecologist, person-environment-behavior systems represent problems to be understood and solved that are simply more important than person-environment-cognition systems or person-environment-attitude systems (p. 19).

The Behavioral Ecology Research Team has emphasized a behavioral perspective in their program of research. The critical issue in the rehabilitative process for a spinal cord injured person is the progress he displays in what he does, how he does it, and where he does it. The period of hospitalization is a period of intensive preparation for life outside the hospital. Thus, not only is it important to assess change in the patient's emerging repertoire of adaptive behaviors, but it is important to assess his emerging accommodation to the environment and his interface with it. Stated in comparative terms, it is more important to measure change in his performance over time than to measure changes in his cognitions over time.

The Behavioral Ecology Research Team uses a nearly-continuous mode of data collection rather than the more traditional mode known as time-sampling, zero-one sampling (Altmann, 1973), or the interval method (Hawkins & Dotson, 1975). The common features of time-sampling are: (a) In each sample period, occurrence, rather than frequency, is reported. (b) Interactions of just a single individual or pair of individuals are recorded. (c) Occurrence usually means "in process" at any time during the sample period. (d) The sample periods are usually short (e.g., 15 seconds), with about 20 sample periods in succession (Altmann, 1973). Time-sampling and continuous recording often yield the same results when the sample periods are consecutive and of short duration. However, McDowell (1973) notes that continuous-recording data can be further subdivided into frequency and duration, permitting analyses of a more molecular nature. He states

> Inasmuch as naturalistic observation studies are often preliminary "fishing expeditions" where data analysis is not fully determined in advance, the method of data collection which offers the greatest post hoc versatility would be most desirable. Therefore, continuous recording would be recommended over time-sampling in preliminary studies of an observational nature (pp. 104-105).

Further, Hawkins and Dotson (1975) argue that the most popular way of checking interval data, the interval-by-interval (I-I) method, is an inadequate assessment of the accuracy and objectivity of the data. They found that I-I reliability scores were (a) highly insensitive as an index of the adequacy of response definitions, (b) a weak measure of observer competence, and (c) insensitive to disagreements between observers or to distortion of experimental effects. Although Hawkins and Dotson offer partial solutions to these problems, they caution that

> ...it is already likely that a significant body of applied behavior analysis has seriously misrepresented to us the relationships between certain environmental factors and certain significant human behaviors ... Further, it is likely that we are ... perpetuating and magnifying our mistakes (p. 376).

The data collection system developed by the Behavioral Ecology Research Team was designed to avoid the shortcomings of the interval method and to enhance post hoc versatility of the data. In this system, observers' descriptions are not broken into intervals and do not include reports of the presence or absence of specified behaviors. Instead, observers produce continuous, time-cued, descriptive narratives by dictating clock time and descriptions of the patient's behavior and situation into hand-held tape recorders. In this way, the data provided by observers are not broken into pre-set intervals, but are as continuous as possible, allowing greater flexibility in the analysis of data.

Another popular method of gathering observational data involves the use of pre-set categories of behavioral events. In this system, observers can easily measure several pre-defined behaviors concurrently, and this technique "circumvents the sometimes difficult task of defining and detecting single units of behavior (as in the case where the experimenter wishes to record ... cooperative play) (Hawkins & Dotson, 1975, p. 359)." Weick (1968) cautions that the behavior-category approach may be troublesome because "it is difficult to know in advance whether the conceptual guidelines are imprecise and whether the translation from concepts to measures has been accurate. If a study proves to be inconclusive, it is difficult to know why (p. 402)." One problem is that the category format is dictated by the assumptions underlying the study. Further, if behavior is categorized as it occurs, time pressures may cause observers to commit errors or omit descriptive detail. The result often is a fragmented impression of the event, without its context. Nonetheless, the observer who watches the event develop is often able to make accurate judgments about the behavior.

A verbal narrative has distinct advantages as a system for recording behavior on the level of molar or complex actions. The vast number of temporal, relational, and linking terms and the even greater number of modifiers and descriptive terms make everyday language invaluable for recording behavior. Allowing observers to describe behavior as it occurs simplifies the problem of translating behavior into pre-defined categories or determining when one behavior ends and another begins. Further, in the narrative system, observers focus on the actual behavior of the target person, and do not have to alternate their attention between that behavior and a list of categories.

Another compelling reason to use the narrative format instead of a category system is that the narrative system permits post hoc coding of the data by independent analysts. A greater variety of categories may be applied in a more leisurely fashion by more coders, and new categories may be developed as needed. Further, the reliability of the coding system may be examined by having several analysts independently code the same narrative observation (Johnson & Bolstad, 1973).

The Behavioral Ecology Research Team uses the narrative system for the reasons above, and because the behavior of the subjects in its program of research often does not readily fit into pre-set categories. As the behavioral repertoire of the spinal cord injured patient grows, and he begins to piece together tiny bits of physical activity into meaningful, purposive behaviors with a discernible beginning and ending, post hoc analysis of narrative data permits a similar piecing together of small bits of data into meaningful clusters.

Separating narrative observation and production of patient records from content coding and data analysis enables the system to be heuristic because (a) new categories of behavior may be added to or deleted from the coding system as needed,

(b) tests of the reliability of the coding system may be conducted repeatedly without interfering with the observation process, and (c) troublesome issues in coding may be discussed and resolved, and old data recoded, without interfering with the ongoing observation process. In short, the separation of data-gathering from data-management ensures the collection of a body of raw data which may be coded, corrected, or recoded, all without irreversible effect on the original narrative observation, and is an important feature of the longitudinal analysis of patient behavior (Willems, 1975, in press).

In this narrative observation system, observers follow a target patient and generate continuous, narrative descriptions of what the patient is doing. A minimum of strictures is placed on the observational process--observations are recorded in the everyday language of the observers. All observations are dictated into battery-operated cassette recorders, and running clock time is noted frequently. A deliberate effort is made by each observer to provide as much relevant detail as possible--the protocol records are rich sources of data on the behavior and behavioral context of patients in the Comprehensive Rehabilitation program at TIRR. The narratives are transcribed and then undergo a three-stage coding process (see Figure 1). The first stage is rough coding, wherein the observer and a coder review the manuscript to make sure that all the necessary details are present for the next coding stage. The second stage, final coding, is done by an independent coder, who marks the narrative protocol off into behavioral units and attaches accompanying coded information to each unit. In the final step, checking, a third coder independently reviews the protocol for accuracy and completeness. Any major disagreements among the three coders must be mutually resolved before the protocol moves to the final stage in data reduction and management--transfer of the coded information to keypunch sheets.

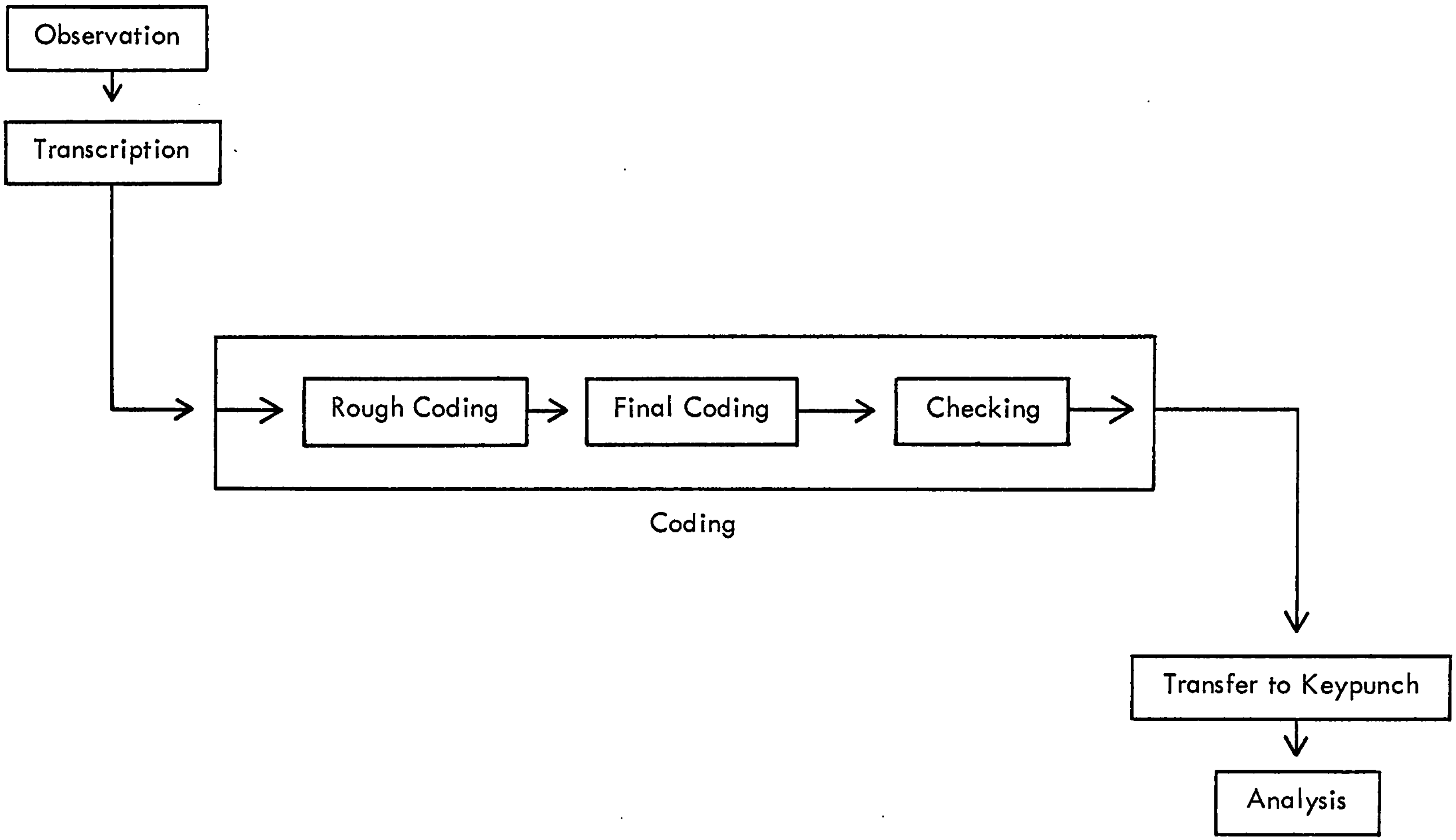


Figure 1. The Data Management System

The basic unit of data in this system, called a chunk, is a molar event in the behavior stream of a patient that can be characterized by a single principal activity that begins at a clearly described starting point, occurs in a characteristic, sustained fashion over time, and ends at a clearly described stopping point. Each chunk is accompanied by the following items of information: (a) location(s) and number of different locations in which the chunk occurred; (b) descriptive label; (c) beginning time; (d) ending time; (e) who else, if anyone, was directly involved in the principal activity of the chunk; (f) who instigated the chunk; and (g) the degree of involvement of the target patient (from active to passive). The format for this information is shown in Figure 2. Marginal entries are attached to reflect mutual verbal interactions between the target patient and other persons (C-codes), and fleeting instances of care or comfort delivered to the target patient by other persons (S-codes), which occur during the chunk. The coding procedures have been standardized (Willems & Crowley, 1976; see Appendix A) and the reliability of both observation and coding are assessed regularly on ten per cent of the data (Dreher, 1975).

The distribution of chunks of patient behavior may be analyzed in many ways. For example, chunks or chunk minutes may be expressed in terms of location, time of day, duration, involvement by others, instigation, or nature of patient involvement. Likewise, locations or settings in the hospital may be characterized in terms of the kinds of chunks or chunk minutes that occur within them, occupancy by time of day, duration, involvement by others, etc. Virtually all of the separate measures may be combined to yield useful, informative compound measures of behavior and its context.

Other combinations are, of course, possible, and many of these combinations have become important conceptual definitions (Willems, in press). The

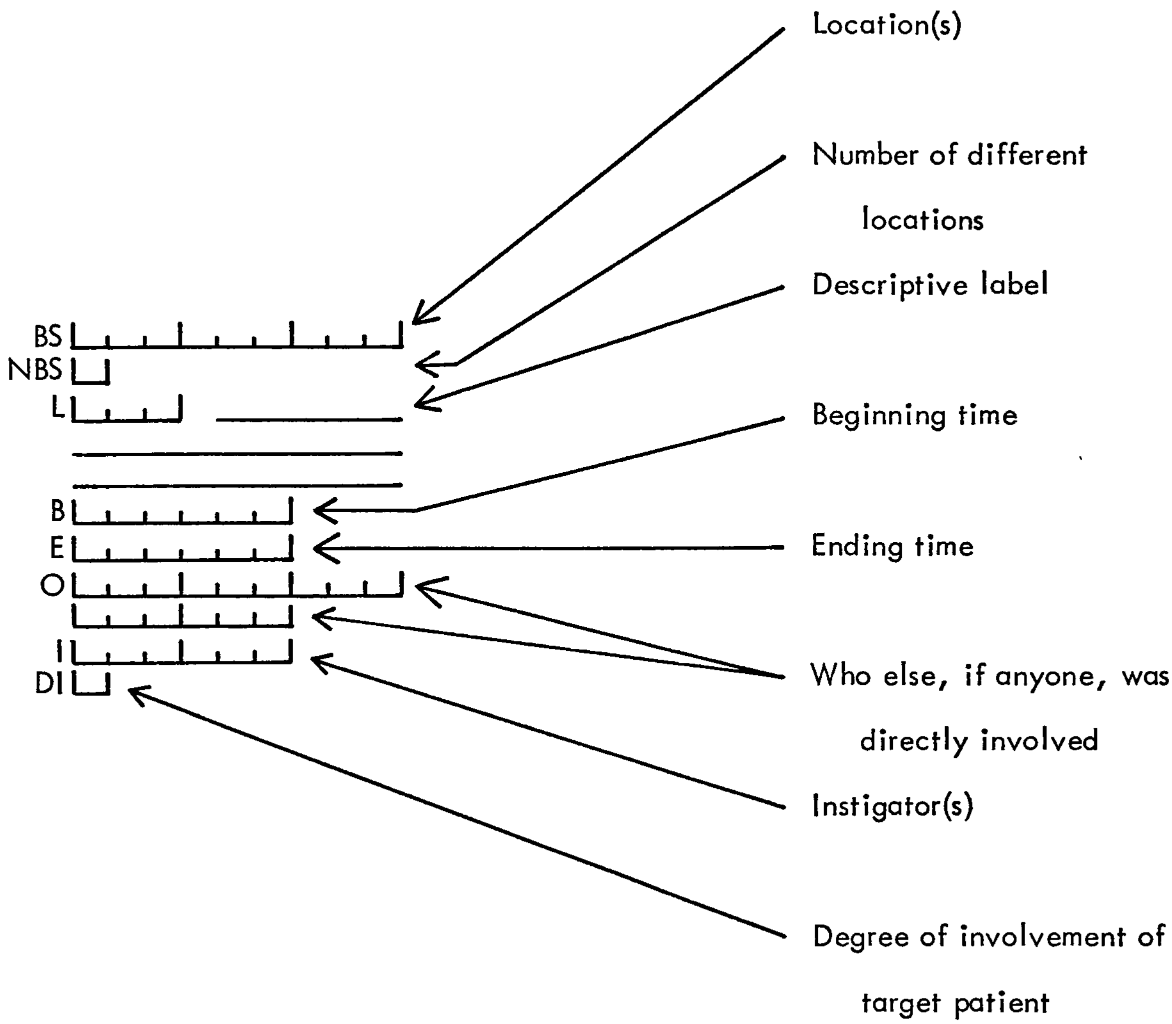


Figure 2. Format for Chunk Code Information

observations yield detailed, quantitative descriptions of repertoires of patient performance; e.g., number of chunks (pace), number of different kinds of chunks (behavioral diversity), number of entries into settings (mobility), number of different kinds of settings entered (environmental diversity), number of instances of overlapping behavior (complexity), rate of self-instigated behaviors, rate of active behaviors, rate of behaviors carried out alone (unaided performance), rate of independent behaviors (self-instigated and unaided), rate of zestful behaviors (self-instigated and active), number of settings in which independence is displayed, and so on. These conceptual definitions, and others, are representative of what is usually the pay-off of the empirical approach to observational research--the identification and selection of a few meaningful measures from a wealth of data points. When such measures are displayed on a regular, cumulative basis (e.g., weekly), they become sensitive and informative indicators of status and change in patient performance. The examples offered below illustrate how this approach monitors such changes.

These illustrations come from the hospital stay of J. H., a patient who was admitted to TIRR in August, 1974, three days after he sustained an accidental injury to his spinal cord (Willems, 1975). Diagnosed as a C 5-6 quadriplegic, J. H. went through the program of comprehensive rehabilitation and was discharged in early December, 1974. Dependent on a wheelchair for ambulation and with only partial return in his hands, J. H. soon returned to his job as a computer programmer-analyst and is now working between 25 and 30 hours a week.

"Independence" is a concept that frequently enters into statements regarding the purposes and philosophy of rehabilitation. In concrete performance terms, independent functioning means behavior that is under the patient's control--he instigates it and performs it alone, without direct involvement of anyone else. Figure 3 displays the weekly rate of such behavioral independence for J. H.

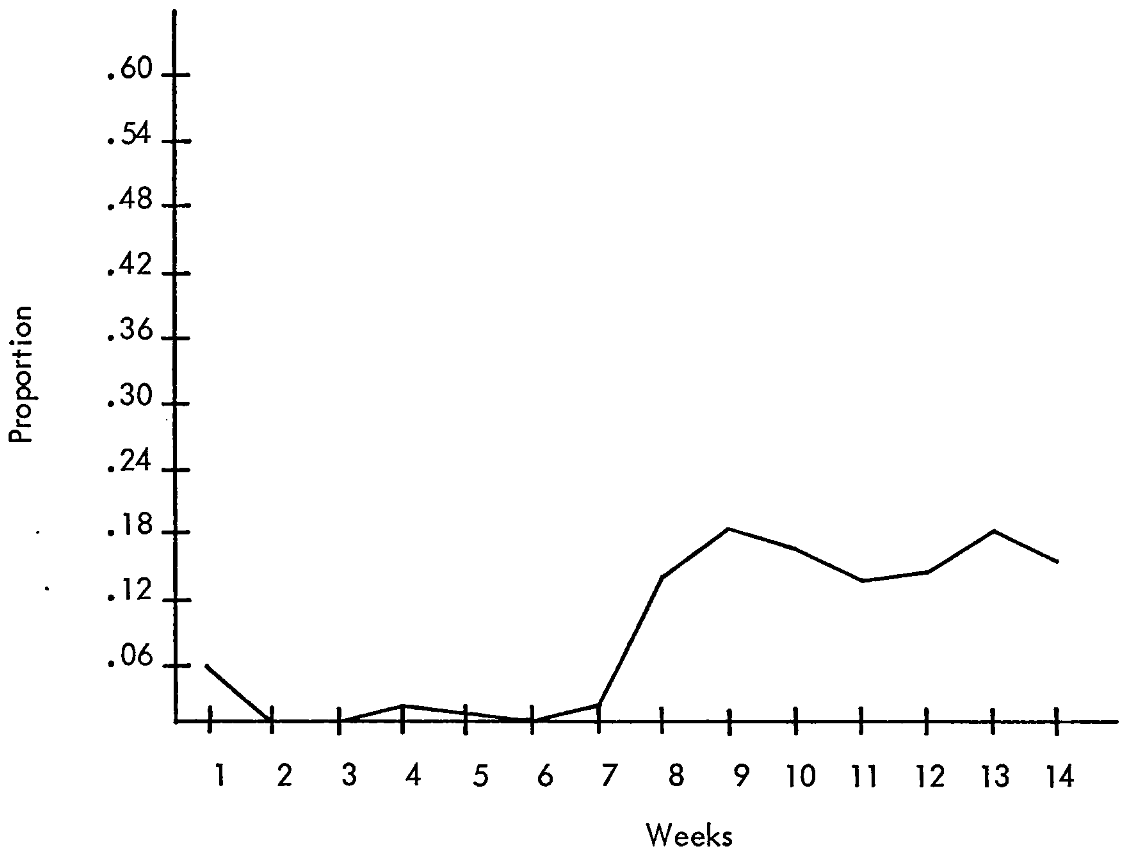


Figure 3. Weekly Rate of Independent Performance:

Proportion of Independent Behaviors (Self-Instigated and Alone) to Total

Performances that are both self-instigated and active are labeled zestful. Calculation of the proportion of such behaviors to total behaviors for each week yields a weekly measure of the patient's rate of behavioral zest. Figure 4 displays the weekly zest rate for J. H. After a six-week period of very low rates, weeks 7 through 10 represented a period of very rapid growth in behavioral zest, followed by a period of leveling off and high variability. Taken together, these measures provide direct, overall, behavioral indicators of the rate and extent to which zestful and independent performance account for J. H.'s total performance.

In addition to such overall rate measures, the observational protocols provide measures of diversity in the patient's repertoire of adaptive skills. Since each coded behavior unit is labeled as to kind of behavior, it is possible to determine the number of different kinds of performances in which the patient engages. For example, Figure 5 displays diversity of active performance, or the number of different kinds of behaviors that J. H. performed actively. On this measure, J. H. showed a diversification of roughly sevenfold in his repertoire of adaptive performances and the phase of most rapid expansion occurred during the two-week period between weeks 6 and 8.

The Behavioral Ecology Research Team participates in treatment teams for their target patients, and presents measures like the ones illustrated above at weekly patient rounds. The clinical personnel are keenly interested in these measures of patient progress, and have provided important feedback on ways to present, illustrate, and describe the data that the Team presents at patient rounds. The most urgent request is that the Team routinize and simplify the data collection and data management systems, because with simplification, routine clinical application will become feasible. The treatment teams are eager to have measures of behavioral progress for all of the spinal cord injured patients. This information has become an important part of planning, assessing, and altering patient programs.

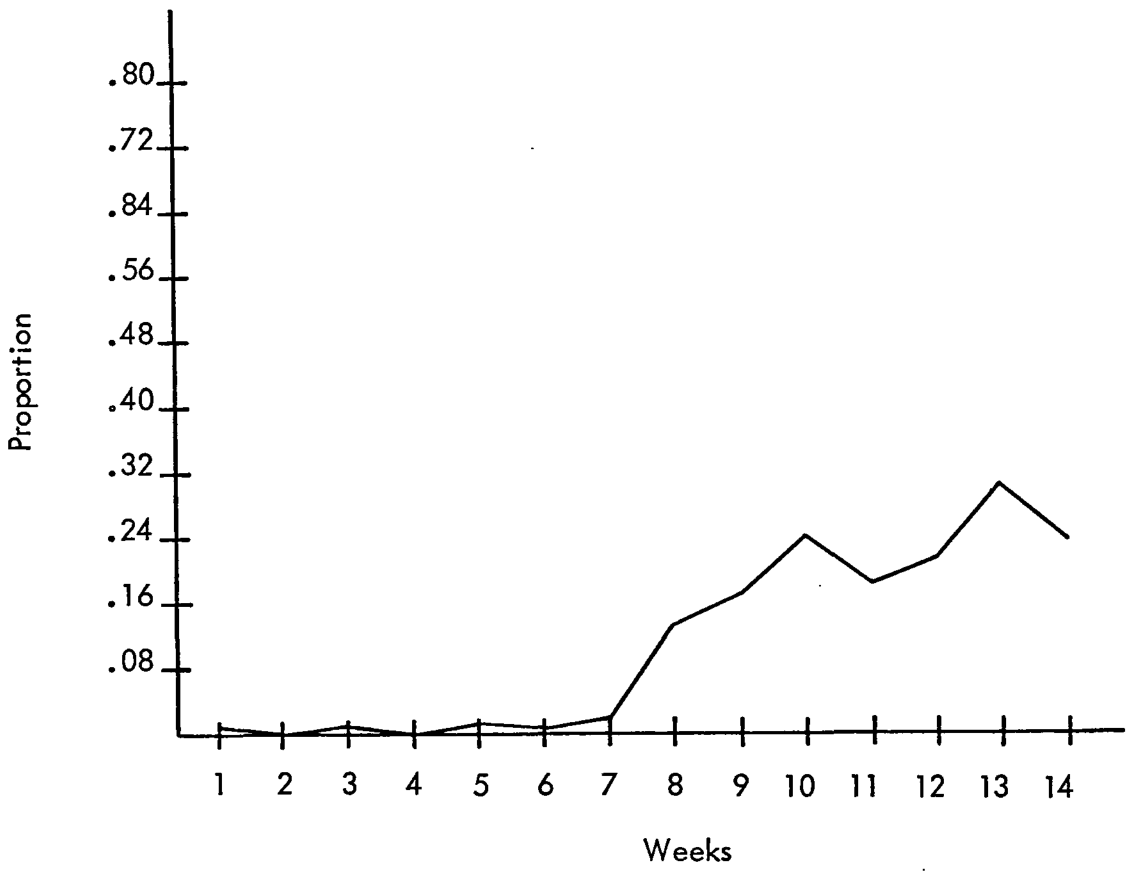


Figure 4. Weekly Rate of Zestful Performance:

Proportion of Zestful Behaviors (Self-Instigated and Active) to Total

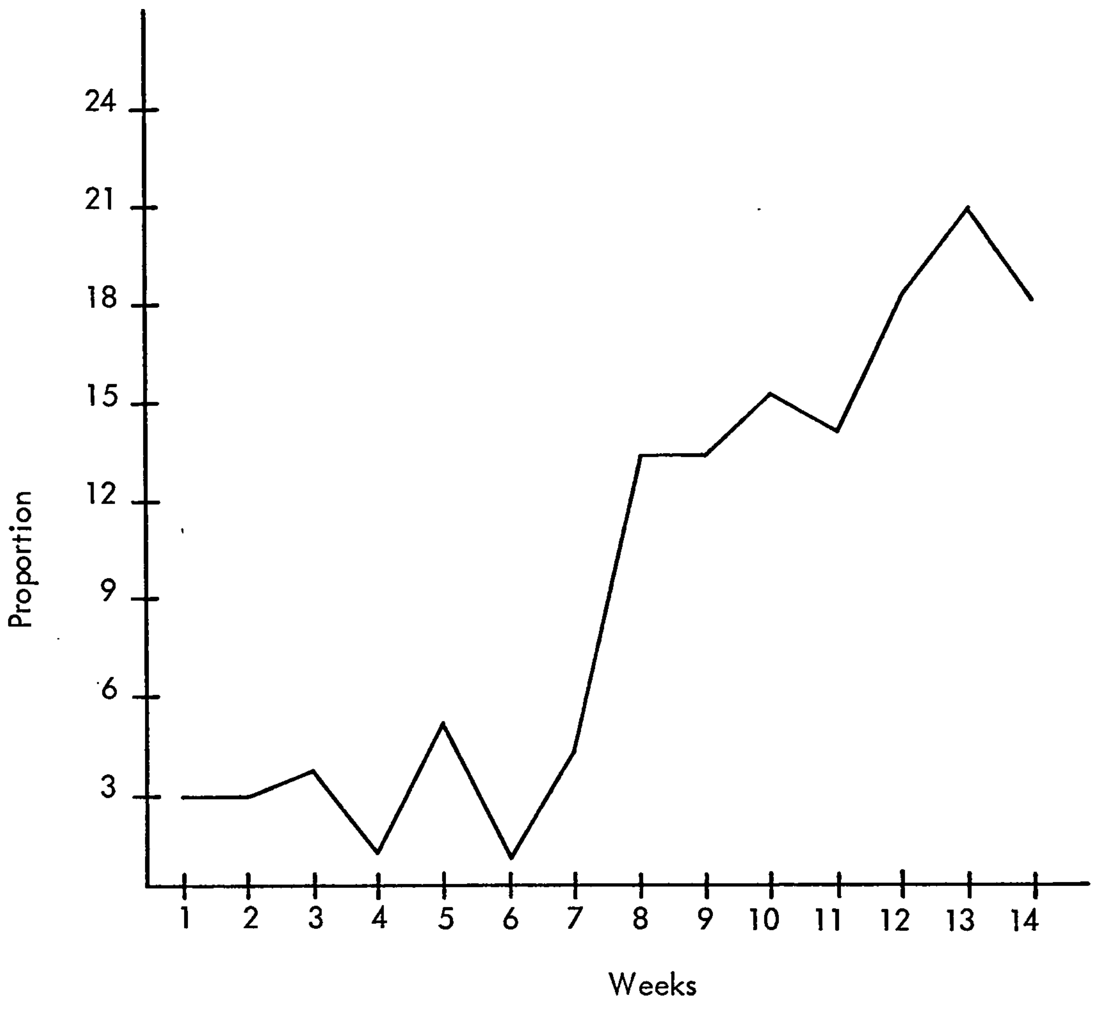


Figure 5. Diversity of Active Performance:

Number of Different Kinds of Active Behaviors

The measures presented above are only three out of 38 that have been developed in the longitudinal analysis of patient behavior. All 38 measures are very sensitive to changes within patients over time. Severe spinal cord injury is catastrophic and, if the person is going to survive and live a meaningful and functional life, then many complex changes must occur during the first year and a half after the injury. One of the best ways to understand that complex process is to describe it as it occurs. Because they describe the gradual re-emergence of behavioral adaptation to the environment following catastrophic loss, data like the examples above provide basic understanding of behavior-environment relations. The data also provide an ongoing set of behavioral indicators of patient progress, and a new information base for staff members to use in making decisions about patient programs.

CHAPTER II

STATEMENT OF THE PROBLEM

Data from the observational studies of 1968 and 1971, cross-sectionally comparing 27 early and advanced (pre-discharge) patients, suggest that, in some important ways, the behavior of patients becomes systematically "normalized" with progress through comprehensive rehabilitation in terms of event structure, diversity, movement, pace, patient initiative, and general level of activity (LeCompte & Willems, 1970; Vineberg & Willems, 1971; Willems, 1972a, 1972b, 1975, in press; Willems, Crowley, & Dreher, 1974). These findings indicate in specific, quantitative, behavioral terms that, in many respects, the hospital was accomplishing much of what it is designed to accomplish. However, there were enough departures and exceptions from the predicted patterns in the cases of several patients to justify a much closer analysis--a longitudinal analysis. The Behavioral Ecology Research Team now begins with patients early in their careers at TIRR (within a week of admission) and monitors each throughout his hospitalization and on into his life beyond the hospital. Studying this relatively long behavioral trajectory within an extended segment of the individual's life in the aftermath of a disabling injury avoids the pitfalls of cross-sectional comparisons of different persons and exploits the strengths of using each person as his own control.

After fully-informed, signed consent has been obtained from the patient, trained observers begin direct observations on a prearranged schedule. Ten 1-1/2 hour observations are spaced across the week in such a way that all of the hours between 7:00 a.m. and 10:00 p.m. are observed during each week. Figure 6 displays an example of one week's observational schedule. The narrative system, with its three-stage coding process described above, is employed to produce patient protocol

	Monday	Tuesday	Wednesday	Thursday	Friday
7:00-8:30 a.m.		X			
8:30-10:00 a.m.				X	
10:00-11:30 a.m.		X			
11:30-1:00 p.m.			X		
1:00-2:30 p.m.					X
2:30-4:00 p.m.			X		
4:00-5:30 p.m.	X				
5:30-7:00 p.m.					X
7:00-8:30 p.m.				X	
8:30-10:00 p.m.	X				

Figure 6. Sample of the Weekly Observation Schedule

records. A member of the Behavioral Ecology Research Team attends weekly patient rounds to provide members of the hospital treatment staff with behavioral measures of patient progress in the rehabilitation program. This information includes cumulative, quantitative descriptions of patient performance such as the measures of patient independence and zest mentioned above. The treatment staff at TIRR have found these measures quite enlightening, and have indicated a desire to receive the most up-to-date information possible, so that they can use the data to assess patient status and to plan patient programs.

The question has been raised whether it is necessary to continue to use the narrative system. Many conceptual definitions have been derived, tested, and found to be valid. Thirty-six patient-subjects have been observed extensively (including eight longitudinally), and the narrative system has reliably disclosed many of the major types and magnitudes of patient behavior as they emerge over time. The research program has reached the point where procedures may now be developed to gather the same information in a less time-consuming, less cumbersome, and less expensive fashion. In the narrative system, approximately six different persons must become involved in the (a) observation, (b) typing, (c) rough coding, (d) final coding, (e) checking, and (f) transferring to computer keypunch sheets. After the initial observation, about 12 person hours are required to process the data. Because of the problem of coordinating persons' schedules, seven to fourteen days are usually required before the data are ready for use, although the period could be shortened if the team were working on only one observation at a time. Research and development notwithstanding, the narrative system requires a formidable amount of time and effort.

A data collection instrument which yields the same data and measures while requiring fewer hours of processing seems desirable at this point. A more efficient system would afford the Behavioral Ecology Research Team three new options: (a) to carry the same number of patients and achieve much faster turn-around time; (b) to carry many more patients at the current turn-around time; or (c) to carry a few more patients and achieve slightly faster turn-around time. Each of these options would provide more up-to-date feedback to the hospital treatment staff for use in assessing the effectiveness of the health care programs they plan and carry out. Furthermore, these developments would provide important steps in the direction of constructing techniques that could be used by clinical personnel.

The purpose of the present research is to develop an alternative system for data collection and management that is more efficient than the narrative system, and that yields the same derived measures of interest. Hopefully, the study will lead to a program of data collection and management that can be used to gather and organize inhospital data on three or four patients with the same investment of time and money that the current system now devotes to a single patient. Savings in person hours alone is not sufficient reason to adopt the new system, however. The distribution of chunks, chunk minutes, and derived behavioral measures must be highly comparable also. This study has four main concerns: (a) construction of a data collection apparatus (tentatively labeled the shortform) and an appropriate coding system; (b) testing the new procedure against the old; (c) assessing the reliability of the new system; and (d) final overall assessment of the new system.

CHAPTER III

METHOD

Construction of Data Collection Apparatus and Coding System

One of the most time-consuming steps in the narrative system is the transcription of the observers' narrative recordings. The shortform circumvents that step by having observers write descriptive notes on pre-set format sheets that are constructed so that observers provide the information necessary for coding patient behavior, while omitting extraneous detail. Each page of the shortform is broken into 15 lines, one line for each minute of observation. Thus, six pages are required to complete every 1-1/2 hour observation period. The shortform also has seven labeled columns in which observers provide, on a minute-by-minute basis, information on (a) time of day, down to seconds; (b) what the target patient is doing; (c) who else, if anyone, is directly involved (identified by staff label and name); (d) who instigates each new activity; (e) the target patient's degree of involvement in the activity (active, passive, or resistive); (f) who, if anyone, converses with the target patient; and (g) location (see Appendix B). A minimum of strictures was placed on the observation process and instructions were deliberately vague. Observers were told to record only "...what the patient did most during the minute," along with the names of persons who became involved in that activity. For example, if a patient ate dinner alone for 50 seconds, a nurse who checked his pulse for 10 seconds should not be recorded as a person directly involved in the patient's eating dinner. If the observer believed that two sequential entries represented the minute better than a single entry, then two entries could be made, but observers were never to resort to listing three or more activities.

A second step in the narrative system that consumes many hours is the rough coding. Rough coding itself requires an average of one hour for every protocol, but the more important problem is that this step requires that two persons be present--the observer and a coder. Two or three days may elapse between observation and rough coding because the protocol must be transcribed and then both observer and coder must arrange to meet at a mutually convenient time. An additional delay may occur if the final coder is not available to begin work when the rough coding has been done. If questions arise in the final coding, the rough and final coders must resolve them mutually. The shortform system combines the rough and final coding steps, so that observer and coder may produce a protocol ready for checking within an hour of the observation. In the present test of the shortform, coding conventions were not prepared beforehand, and instructions for coders were deliberately vague. Coders were asked to "Mark off the patient behavior into chunks, and provide all necessary accompanying codes." Throughout the development of the prototypical shortform and coding system, observers and coders were asked for feedback, comments, and advice for clarifying or simplifying the new system.

Testing the New Procedure Against the Old: Convergent Validity

In this phase of evaluation, ten pairs of observers, one observer using the narrative mode and the other using the shortform mode, observed a single target patient simultaneously for 1-1/2 hours each. The target patient was a subject in the longitudinal study. Sample sections from one of the paired observations are shown in Figure 7 (narrative observation) and Figure 8 (shortform observation). The paired observations spanned the regular data-collection period between 7:00 a.m. and 10:00 p.m. In every case, the shortform observation and coding were performed independently of the ongoing narrative system. The observation periods used for this analysis were among those already scheduled in the longitudinal study. The

7:29:50	NA Howard comes up to Anderson's bedside with a blood pressure machine.
7:30:34	Howard says something to Anderson and he responds. She touches the thermometer to his lip but he won't take it.
7:30:45	She is still trying to get the thermometer into his mouth, but he is shaking his head, no.
7:30:55	He says his nose is stuffed up. She is holding the thermometer over his mouth, not putting it in.
7:31:05	Howard puts the thermometer down and picks up the blood pressure (bp) cuff.
7:31:23	Howard begins wrapping the bp cuff around Anderson's arm. They don't converse.
7:31:48	Howard inflates the bp cuff.
7:32:07	She takes the bp cuff off. Anderson is apparently asleep. There is no conversation.
7:32:43	Howard begins taking Anderson's pulse.
7:33:05	She stops doing that. Anderson is still lying quietly with his eyes closed.
7:33:40	Howard leaves the area, leaving Anderson alone.
7:34	Howard returns and begins to close the curtains around Anderson.
7:34:05	The curtains are closed and Howard is in there with Anderson.

Figure 7. Sample of Narrative Observation

Time	Activity	Others	Instig.	DI	C	Location
7:29	Lying in bed	none	none	—	-θ-	Ward
7:30	Lying in bed	none	none	—	-θ-	"
7:31	Having bloodpressure taken	NA Howard	NA Howard	passive (2)	-θ-	"
7:32	Lying in bed	none	none	—	-θ-	"
7:33	Having pulse taken Lying in bed	NA Howard none	NA H. none	2 —	-θ-	"
7:34	Curtains closed around him — Patient Care	NA Howard	NA Howard	2	-θ-	"

Figure 8. Sample of Shortform Observation

data obtained from the regular narrative system were transferred to keypunch sheets and became part of the feedback offered during patient rounds. Those data also served as the standard against which the shortform data were assessed. Seven observers and two coders from the narrative system volunteered to participate as observers and coders in the shortform system. Each observer was given approximately five minutes of verbal instruction on the use of the shortform, and then the observers practiced with the form until they felt accustomed to the location of the various columns on the form. These practice sessions lasted between five and fifteen minutes. The coders were given verbal instructions only; no practice materials were provided. Extensive or rigorous training of observers and coders was not deemed necessary in light of the expertise each had developed in the narrative system. It was hoped that the two systems would be so similar that observation and coding skills developed in one system would transfer quickly and easily to the other.

Congruence: 5-Minute Segments

In this analysis, the two distributions of patient activities (chunks) across the 1-1/2 hour observation were compared in terms of the five-minute segment analysis developed by Dreher (1975). The guiding question here was whether the two descriptions of patient behavior were reasonably congruent or whether the accounts seemed divergent or systematically different. In other words, to what extent did the two systems map behavior in the same way? The observations in each pair were broken into 18 5-minute segments. Each pair of segments was compared in terms of the type (or label) of chunks occurring within them. Because some chunks began and/or ended within the segments, and because chunks were coded for beginning time and ending time, it was possible to calculate the exact number of seconds for which there was complete agreement or disagreement about what kind of patient activity had occurred.

For example, consider a hypothetical 5-minute segment between 1:00 and 1:05. If in one version, the patient was coded as eating lunch from 1:00:00 until 1:02:30 and then coded as wheeling himself out of the cafeteria from 1:02:30 to 1:05:00, but in the other version the patient was coded as eating from 1:00:00 until 1:02:45 and then as transporting from 1:02:45 until 1:05:00, there were 285 seconds of agreement and 15 seconds of disagreement in that segment. The number of seconds of disagreement for each 5-minute segment was tallied and a mean agreement score for the 18 segments in each paired observation was compiled, as well as a grand mean for the 10 paired observations. Disagreements were compiled and categorized for further analysis (e.g., disagreements due to differences in the observational accounts, differences in the coding based on coder interpretation of observer description, or differences between the narrative observation and shortform systems).

General Information Provided in Chunks and Accompanying Codes

Chunks and chunk codes from each pair of observations were compared using the parallel chunk procedure developed by Dreher (1975). The focus here was on those portions of the patient's behavior stream for which there was parallelism, or high agreement about what the patient was actually doing. For example, consider the following lists of chunks:

Version 1	Version 2
Eating lunch	Eating lunch
Transporting to ward	Transporting to ward
Transferring to bed	Transferring to bed
Lying in bed	Reading book
Reading book	

If these lists represented the data from a paired observation, then "Eating lunch" in

Version 1 and "Eating lunch" in Version 2 were considered parallel chunks because they (a) agreed with respect to the type of patient activity and (b) occurred in the same relative place in the patient's behavior stream. Similarly, "Transporting to ward," "Transferring to bed," and "Reading book" were considered parallel chunks. "Lying in bed" in Version 1 has no counterpart in Version 2, and thus would be dropped from this analysis.

Wherever a pair of parallel chunks occurred, the pair of chunks was examined for agreement or disagreement on (a) location; (b) number of different locations; (c) principal activity of the chunk; (d) beginning time; (e) ending time; (f) elapsed time; (g) others, if any, who were directly involved in the activity; (h) instigator(s); (i) degree of involvement by the target patient; and (j) persons, if any, who conversed with the patient during the chunk. The number of agreements in each of these ten chunk codes was tallied, and the ratio of agreements in each of the chunk codes to the total number of parallel chunks was calculated using the following formula:

$$X = \frac{2\ (\text{number of chunks on which the narrative and shortform agree})}{(\text{number of chunks claimed by the narrative} + \text{number of chunks claimed by the shortform})}$$

Disagreements were compiled and categorized for further analysis, in an attempt to determine whether the disagreements were due to differences in the underlying observational data, differences in coder interpretation of observational description, or differences between the narrative and shortform systems.

Differences in the Distribution of Time into Chunks

Due to the exigencies of coding in both the narrative and shortform systems, differences in beginning time, ending time, and elapsed time were to be expected. The narrative allowed observers to provide exact time entries for the coders to use for time (e.g., 1:02:47), but the shortform required coders to use approximate time entries (the midpoint between whole minutes was the only fraction allowed). In the narrative system, a chunk could begin or end at any second, but in the shortform system, chunks could begin or end only on the minute or half-minute. Disagreements in data on time (beginning time, ending time, and elapsed time of chunks), which were identified in the analyses of congruence and parallel chunks, were examined here by comparing the paired distributions of chunk seconds. An attempt was made to identify (a) the disagreement due to substantive differences in the data provided by observers, (b) disagreement due to substantive differences in the coding of the observational records, and (c) disagreement due only to differences between the narrative and shortform systems. Disagreement due to differences in the two systems was defined as disagreement in time at the junction between pairs of parallel chunks.

Comparison of Derived Measures

The derived measures (e.g., patient-instigated chunks, active chunks, independent chunks, zestful chunks, nonidle chunks, and idle chunks) produced in each pair of protocols were compared. The measures derived from the narrative system were treated as the standard against which the shortform data were assessed. Agreement scores for each of the derived measures were calculated as follows:

$$X = \frac{2\ (\text{number of chunks on which the narrative and shortform agree})}{(\text{number of chunks claimed by the narrative} + \text{number of chunks claimed by the shortform})}$$

Differences revealed in the analyses of congruence and parallel chunks were expected to affect the derived measures; therefore, the differences in derived measures were examined in hopes of discerning the source of discrepancy (i.e., differences in underlying observational data, differences in the coding of that data, or differences between the narrative and shortform systems).

Reliability of the Shortform

For this evaluation, ten different pairs of observers observed a single target patient simultaneously for 1-1/2 hours each, in a new series of observations. Both observers in each pair used the shortform. The paired observations spanned the hours between 7:00 a.m. and 10:00 p.m. Five observers and two coders volunteered to participate in this phase of the shortform evaluation. The observers were paired so that each possible pair completed an observation and each observer completed four observations. The coders were assigned so that each coded twice with a single observer. The observer-coder teams are illustrated in Figure 9. Data collection for this phase of the shortform study was accomplished after all of the narrative/shortform pairings had been completed. No further training was offered to the observers or coders. The last coding step--checking--was deleted from the data management system because only two coders participated in this phase of the study. The coders were urged to work independently and to check their own work closely. Once collected and coded, the data were assessed from four perspectives.

Congruence: 5-Minute Segments

The pairs of protocols were divided into 5-minute segments and compared as described above. Disagreements were compiled and categorized for further analysis (e.g., disagreement due to differences in data provided by observers, or differences in coder interpretation).

General Information Provided in Chunks and Accompanying Codes

Parallel pairs of chunks and their accompanying codes were compared as described above. Disagreements were compiled and categorized for further analysis

Observer Pool

1 2 3 4 5

Coder Pool

A B

Paired Observation	Team 1		Team 2	
1	1	A	2	B
2	1	A	3	B
3	1	B	4	A
4	1	B	5	A
5	2	A	3	B
6	2	A	4	B
7	2	B	5	A
8	3	A	4	B
9	3	A	5	B
10	4	A	5	B

Figure 9. Teams of Observers and Coders
for the Ten Paired Shortform Observations

(e.g., disagreement due to differences in the observation records, or in the coders' interpretations of those records).

Fine-Grained Level of Agreement

Every pair of line entries in the minute-by-minute data provided on the short-form was examined for agreement. The criterion for agreement between a pair of line entries was that every pair of cells within those line entries be identical; i.e., that location, activity, other(s), instigation, degree of involvement, and conversation all agree. The number of line entries with perfect agreement for each paired observation was tallied and the percent agreement was calculated for each observation. In addition, a mean agreement score for line entries in the ten paired observations was calculated. Agreement scores were also calculated for the separate columns within each paired observation (e.g., location, activity, etc). Mean agreement scores were calculated for each of the columns in the ten paired observations. Disagreements were compiled and categorized for further analysis (e.g., errors of omission or mistaken identity).

Comparison of Derived Measures

The derived measures (patient-instigated chunks, active chunks, independent chunks, zestful chunks, nonidle chunks, and idle chunks) produced in the paired observations were compared, and agreement scores were calculated as described above. Disagreements were compiled and categorized for further analysis (e.g., disagreement due to difference in observation records, or in coder interpretation of observer description).

CHAPTER IV

RESULTS

Comparison of Narrative and Shortform Data

Congruence

Table 1 displays the mean agreement scores for 5-minute segments in the paired narrative/shortform observations. Comparison of 5-minute segments within each paired 1-1/2 hour observation resulted in agreement scores ranging from 64.7 to 99.6 per cent, and a grand mean agreement score of 87.8 per cent (Table 1, first column of per cents). Eight of the ten agreement scores exceeded 85 per cent. The discrepancies between narrative and shortform data fell into three major categories: (a) differences in chunk labels based on coder interpretation of observer description; (b) differences due to approximated beginning and ending times in the shortform; and (c) differences attributable to the problem of observing and coding the specific activity, "Watching TV." All other discrepancies seemed more nonsystematic. Removing the effect of the activity, "Watching TV" (by considering whatever other activities the patient was engaged in), raised each of the three affected agreement scores and raised the grand mean to 91.5 per cent (Table 1, second column of per cents).

The number of chunks per observation is sometimes thought to be related to agreement score (Dreher, 1975): the fewer the chunks, the higher the agreement score. Table 2 lists the ten paired observations, ranked in terms of the number of parallel chunks identified within them, and the 5-minute segment agreement scores for each paired observation. Although the observation with the fewest parallel chunks (3) did have the highest agreement score (99.6 per cent), and the observation

Table 1

Agreement Scores (5-Minute Segments):

Narrative/Shortform Comparison

Paired Observation Number	Mean Agreement Score (With TV)	Mean Agreement Score (Without TV)
1	95.4	no change
2	88.6	no change
3	99.6	no change
4	89.2	no change
5	87.5	no change
6	92.6	no change
7	79.7	83.5
8	85.6	92.9
9	64.7	91.2
10	94.9	no change
Mean	87.8	91.5

Table 2

Number of Parallel Chunks and Agreement Scores (5-Minute Segments): Narrative/Shortform Comparison

Paired Observation Number	Number of Parallel Chunks	Mean Agreement Score (With TV)	Mean Agreement Score (Without TV)
3	3	99.6	99.6
4	6	89.2	89.2
2	7	88.6	88.6
6	11	92.6	92.6
1	12	95.4	95.4
8	13	85.6	92.9
7	14	79.7	83.5
10	16	94.9	94.9
9	18	64.7	91.2
5	27	87.5	87.5

with the most chunks (27) had one of the lower agreement scores (87.5 per cent), the relationship between number of parallel chunks and agreement scores was not strong. The Spearman rank order correlation between the number of parallel chunks and 5-minute segment agreement score (including the coding of "Watching TV") was -.576, and not significant for N-2=8 degrees of freedom, α=.01. Removing the effect of TV chunks did not improve the relationship between number of parallel chunks and agreement score (r_s = -.345, not significant for df=8, α=.01).

Chunks and Accompanying Codes

Table 3 displays the agreement scores on parallel chunks and accompanying chunk codes for the ten paired observations. The first row below the paired observation number lists the number of parallel chunks identified. The remaining rows list the number of parallel chunks for which there was exact agreement on: (a) location; (b) number of different locations; (c) others, if any, who were involved in the activity; (d) instigator(s) of the activity; (e) degree of involvement of the target patient; (f) others, if any, who conversed with the target patient during the activity; (g) beginning time; and (h) ending time.

The number of parallel chunks in the paired observations ranged from 3 to 27, with a mean of 13 (see Table 3). The total number of parallel chunks for the ten observations was 127. Agreement on location and number of different locations was 100 per cent across every parallel chunk in the pairs of observations. Agreement on the coding of others (if any) who were directly involved in the activity was 100 per cent for 7 of the 10 pairs of observations, and mean agreement for this category was 96.9 per cent. Agreement on instigation had a mean of 93.7 per cent across the 10 pairs of observations, and mean agreement on the patient's degree of involvement was 95.3 per cent.

Table 3

Agreement Scores on Chunks and Accompanying Codes:

Narrative/Shortform Comparison

	Paired Observation Number										Total	
	1	2	3	4	5	6	7	8	9	10	#	%
Number of Parallel Chunks	12	7	3	6	27	11	14	13	18	16	127	
Agreement on (#Chunks):												
Location	12	7	3	6	27	11	14	13	18	16	127	100.0
Number of Locations	12	7	3	6	27	11	14	13	18	16	127	100.0
Others	12	7	3	6	25	11	14	12	17	16	123	96.9
Instigators	12	6	3	6	26	10	12	12	18	14	119	93.7
Degree of Involvement	11	6	3	6	25	10	14	13	18	15	121	95.3
Conversation	10	6	2	5	15	9	8	12	11	8	86	67.7
Beginning Time*	1	2	1	1	1	1	1	1	2	1	12	9.4
Ending Time*	1	1	1	1	1	1	1	2	2	1	12	9.4

*Includes the beginning time of first chunks and the ending time of last chunks in each protocol, which are governed by convention (Willems & Crowley, 1976).

Agreement on the coding of persons (if any) who conversed with the patient during the 127 parallel chunks was 67.6 per cent, and was considerably lower than the agreement scores mentioned above. Other investigators have discovered the same phenomenon in their own work; i.e., inter-observer agreement on conversational behaviors is lower than interobserver agreement on other behaviors (Bijou, Peterson, & Ault, 1968; McDowell, 1973). For example, McDowell (1973) studied caretaker-infant interactions by having observers record five caretaker behaviors (holds, feeds, diapers, talks to infant, caretaking) and five infant behaviors (asleep, bottle in mouth, eats, vocalizes, cries). Interobserver agreement over all ten behaviors had a range of .63 to .99, with a median of .93 for a continuous recording mode. The lowest interobserver agreement scores were for the behaviors "talks to infant" and "vocalizes"--both were .63. Bijou, Peterson and Ault (1968) studied the behavior of a single child in a class of nursery school children. Interobserver agreement on social contacts (verbal and physical) exceeded .82 while agreement on sustained activities (school behavior) exceeded .95. However, Bijou, Peterson and Ault do not report an interobserver agreement score for verbal social contacts alone. Hawkins and Dotson (1975) suggest that a major source of error in obtaining accurate and objective data is that the behavior in question may be difficult to detect because of its subtlety or complexity. In this context, the lower agreement scores obtained for conversation in the present study seem reasonable, particularly because the criterion for coding conversation is the observer's report of a mutual verbal interaction. In other words, an observer must report a rather subtle behavior on the part of both the target patient and another person (a more stringent criterion for coding verbal behavior than that reported by either McDowell (1973) or Bijou, Peterson, and Ault (1968)). The lack of agreement in this study was largely due to the number of instances in which one observer reported mutual conversation and the other did not, rather than instances in which observers disagreed on the identity of a particular person.

Finally, agreement on both beginning time and ending time of patient activities was 9.4 per cent, and reflects (a) application of the coding conventions used to define the beginning time of first chunks and ending time of last chunks in each observational protocol (Willems & Crowley, 1976), and (b) the rare occasions in which patient behavior began or ended precisely on the minute or half-minute. The first chunk in every narrative protocol begins when the narrative observation itself begins, by convention (Willems & Crowley, 1976). Similarly, the last chunk ends when the observation ends. These conventions were used in the coding of the shortform also. Therefore, the first chunks in paired observations had identical beginning times and the last chunks had identical ending times. Further, the narrative allowed observers to provide exact time entries for the coders to use for beginning time or ending time on all other chunks (e.g., 11:08:23), but the shortform required coders to use only whole minutes (e.g., 11:08:00) or half-minutes (e.g., 11:08:30) for beginning time or ending time. Therefore, agreement on time codes for any pair of chunks (except first and last chunks) was rare, and occurred only when the narrative chunks actually began or ended on a whole minute or half-minute. Low agreement on time codes was expected; however, it is important to examine the time data closely, as they represent the greatest potential source of systematic disagreement between the narrative and shortform systems. The next section presents an analysis of the disagreement on distribution of time in the paired observations.

Distribution of Time into Chunks

Table 4 displays the ten paired observations in terms of the number of seconds of disagreement found within them. The first row below the paired observation

Table 4

Distribution of Seconds of Disagreement:

Narrative/Shortform Comparison

	Paired Observation Number										Total	
	1	2	3	4	5	6	7	8	9	10	#	%
Total Seconds of Disagreement	246	615	19	584	677	402	1098	779	1904	276	6600	
Time Convention	115	78	19	105	201	94	59	138	168	131	1108	16.79
Chunk on Form, S on Narrative	120	30			30		831		175		1186	17.97
Gap	11				9	13					33	.50
Chunk on One, Nothing on Other				31				102		75	208	3.15
TV-Related							208	363	1494		2065	31.29
Lying-Related		442		448	319	277		176	67	70	1799	27.26
Exercise-Related					118						118	1.79
Miscellaneous		65				18					83	1.26

number shows the total seconds of disagreement for the observations. The following rows list the number of seconds within several classes of disagreement. The last two columns show the total and per cent for each row.

There were 6600 seconds of disagreement out of 54000 seconds of patient observation (12.22 per cent). Three classes of disagreement were directly attributable to procedural differences between the coding systems. These classes were labeled (a) Time Convention; (b) Chunk on Form, S on Narrative; and (c) Gap. The other classes of disagreement were routine problems in the narrative system and therefore were not strictly attributable to procedural differences between the narrative and shortform systems. The latter classes accounted for 4273, or 64.74 per cent of the 6600 seconds of disagreement.

Time convention differences occurred when parallel chunks seemed to begin or end at slightly different times because the observer or coder of the shortform had rounded to the minute or half-minute (e.g., the narrative beginning time was 9:06:18 and the shortform beginning time was 9:06:30). Discrepancies of this type were usually 14 seconds or less, and occurred at the junction between parallel chunks. Time convention differences accounted for 1108, or 16.79 per cent, of the disagreement.

The narrative system had a coding option that the shortform did not--the S-code, which covered short, fleeting instances of care or comfort delivered to the patient by someone else. When an activity involving the patient and another person was highly intermittent or sporadic, the coder in the narrative system could elect to code the activity with an S or series of Ss attached to the major units, or chunks. The shortform observer and coder did not have this option; they had to either code the activity as a chunk or omit it. This sort of discrepancy--chunk on shortform, S on narrative--accounted for 1186, or 17.97 per cent, of the 6600 seconds of disagreement.

The third class of time differences was labeled, "Gap." Coders in the narrative system could elect to leave short periods of patient behavior uncoded if the activity was not related to the chunks that preceded or followed it. Such periods never exceeded 14 seconds, because any episode 15 seconds or longer had to be coded as a chunk, by convention (Willems & Crowley, 1976). The shortform system did not allow for gaps, but absorbed them along with any other fleeting occurrences in its summary approach to patient behavior. Gaps accounted for only 33, or .5 per cent, of the 6600 seconds of disagreement.

The remaining classes of disagreement were not directly attributable to procedural differences between the narrative and shortform systems. These disagreements were due to substantive differences in the descriptions provided by observers, and/or to differences in the interpretations drawn by coders. The class, "Chunk on One, Nothing on Other," covered all non-parallel chunks in the paired observations, except nonparallel "Lying," "Sitting," and "Watching TV" chunks, which were examined separately. "Chunk on one, nothing on other" occurred (a) when an activity was described in enough detail by one observer that the activity was coded as a chunk, but not mentioned at all by the other observer; or (b) when one coder included a specific activity in an ongoing chunk, while the other coder coded the activity as a separate or overlapping chunk. This type of disagreement accounted for 208, or 3.15 per cent, of the 6600 seconds of disagreement.

Disagreement related to the coding of the activity , "Watching TV," occurred in only three of the paired observations, but accounted for the bulk of the disagreement within those observations. This type of disagreement accounted for 2065, or 31.29 per cent, of the 6600 seconds of disagreement. The problem of coding this specific activity will be discussed below.

Disagreements related to the coding of "Lying" occurred in seven of the ten paired observations. The coding of "Lying" is governed by convention (Willems & Crowley, 1976) and the beginning time and ending time of a "Lying" chunk denote a period in which the patient was idle and horizontal. ("Sitting" refers to the condition of being idle and vertical.) Most disagreements occurred (a) when one coder elected to interrupt a "Lying" chunk with a short, non-idle chunk and the other coder did not, or (b) when one coder began or ended a "Lying" chunk before the other coder did. Disagreements in the coding of "Lying" also affected the derived measures of independence, zest, non-idle behavior, and idle behavior, and accounted for 1799, or 27.26 per cent, of the total seconds of disagreement.

Disagreements related to the coding of "Exercising" occurred in only one of the ten paired observations. These disagreements occurred (a) when one coder labeled an activity "Exercising Arm," while the other coder called it, "Exercising Wrist," and (b) when one coder included an idle episode ("resting") in an "Exercising" chunk, while the other coder concluded that the episode was not related to the exercise, and called it "Lying." These differences accounted for 118, or 1.79 per cent, of the 6600 seconds of disagreement.

The last category--Miscellaneous--contains two types of disagreement. In one case, one pair of parallel chunks in the middle of a series of parallel chunks began 65 seconds apart. The most plausible explanation is that one observer somehow "lost" a minute of time at the beginning of the chunk and caught up again before the end of the chunk. In the second case, one coder identified a "Drinking" chunk where the other identified a "Wiping Face" chunk. These discrepancies were singular occurrences and accounted for 83, or 1.26 per cent, of the 6600 seconds of disagreement.

Derived Measures

Agreement on the measures for patient instigation, active involvement, independence, zest, non-idle behavior, two idle behaviors ("Sitting" and "Lying"), and one semi-idle behavior ("Watching TV") was calculated for the pairs of observations. Agreement on the same derived measures was also calculated for the pairs of parallel chunks in the paired observations. Table 5 lists the overall percent agreement scores for derived measures from all of the chunks in the paired observations. Table 6 lists the percent agreement scores for derived measures from parallel chunks only.

There was generally high agreement in the derived measures obtained in the all-chunk analysis. The derived measures compiled from parallel chunks displayed slightly higher agreement than measures from all chunks, but this was expected because nonparallel chunks had been deleted. Every agreement score of "0" in the parallel chunk analysis reflected the situation where one system yielded one chunk and the other system yielded none; e.g., one system identified one patient-instigated chunk and the other system yielded no patient-instigated chunks.

From Table 5 (All Chunks), it can be seen that agreement on patient-instigation was 100 per cent for five pairs of protocols (no patient-instigated chunks were identified). Two pairs of protocols yielded 0 per cent agreement (in one case the narrative yielded one patient-instigated chunk and the shortform yielded none; in the other case, the situation was reversed). Three pairs of protocols yielded agreement scores of 75, 73, and 67 per cent. The first two scores reflected differences of two and three patient-instigated chunks, respectively, and in each case the "surplus" chunks were "Watching TV." The agreement score of 67 was caused by one chunk identified in the shortform system that was not so identified in the narrative system.

Table 5

Percent Agreement on Derived Measures from All Chunks:

Narrative/Shortform Comparison

Percent Agreement on Number of:	Paired Observation Number									
	1	2	3	4	5	6	7	8	9	10
Patient-Instigated Chunks	100	100	100	100	100	0	0	75	73	67
Patient-Active Chunks	100	0	100	100	80	0	100	100	100	100
Independent Chunks	100	100	100	100	100	100	0	67	73	100
Zestful Chunks	100	100	100	100	100	0	100	100	100	100
Non-Idle Chunks	86	100	100	91	93	88	76	92	84	100
"Lying in Bed" Chunks	92	86	100	88	96	77	77	75	95	93
"Sitting" Chunks	100	100	100	100	100	100	100	100	100	100
"Watching TV" Chunks	100	100	100	100	100	100	67	75	67	100

Table 6

Percent Agreement on Derived Measures from Parallel Chunks:

Narrative/Shortform Comparison

Percent Agreement on Number of:	Paired Observation Number									
	1	2	3	4	5	6	7	8	9	10
Patient-Instigated Chunks	100	100	100	100	100	0	0	75	100	0
Patient-Active Chunks	100	0	100	100	80	0	100	100	100	100
Independent Chunks	100	100	100	100	100	100	0	67	100	100
Zestful Chunks	100	100	100	100	100	0	100	100	100	100
Non-Idle Chunks	100	100	100	100	100	100	100	92	100	100
"Lying in Bed" Chunks	100	100	100	100	100	100	100	100	100	100
"Sitting" Chunks	100	100	100	100	100	100	100	100	100	100
"Watching TV" Chunks	100	100	100	100	100	100	100	75	100	100

In Table 6 (Parallel Chunks), the agreement scores for patient-instigated chunks remained the same for eight of the ten pairs of protocols. One score improved to 100 per cent as unmatched chunks were deleted; however, one score fell from 67 per cent to 0 per cent. This occurred because, although each system identified one patient-instigated chunk, they identified different chunks, chunks which were not paired together.

The trend toward high agreement held for the rest of the derived measures, and agreement remained the same or improved when nonparallel chunks were removed. Agreement on active involvement was 100 per cent in seven pairs of protocols in terms of both all and parallel chunks. Agreement on independent patient chunks (self-instigated and nobody else involved) was 100 per cent in seven of the all-chunk comparisons, and eight of the parallel chunk comparisons. Agreement on patient zest (self-instigated and active) was 100 per cent for nine of the ten paired protocols in terms of both all chunks and parallel chunks.

Measures on non-idle, idle, and semi-idle behaviors displayed slightly more variability than the measures above. From Table 5, it can be seen that agreement on non-idle chunks ranged from 76 to 100 per cent, with only three pairings yielding agreement scores of 100 per cent. Deleting nonparallel or unmatched chunks (Table 6) yielded nine agreement scores of 100 per cent. Similarly, the agreement scores for "Lying in Bed" chunks ranged from 75 to 100 per cent in terms of all chunks, with only one pairing yielding an agreement score of 100 per cent. Removing unmatched chunks yielded perfect agreement. There was also perfect agreement on "Sitting" chunks in terms of all chunks and parallel chunks (none were identified in any of the paired observations).

"Watching TV" chunks seemed to present problems--both in terms of identifying the chunks themselves and in terms of the measures of patient instigation and patient independence. TV chunks occurred in only three paired observations (Paired Observation Numbers 7, 8, and 9). In terms of all chunks, the agreement on number of "Watching TV" chunks was 67, 75, and 67 per cent in the three observations where TV watching occurred. Deletion of nonparallel chunks improved the scores to 100, 75, and 100 per cent, respectively. "Watching TV" is coded by convention (Willems & Crowley, 1976) as patient-instigated and performed alone, unless the observer makes it abundantly clear that the activity is not self-instigated or performed alone. Disagreements in coding the activity of "Watching TV" may result in disagreements in measures of patient instigation and patient independence (self-instigated and performed alone)--and this was the case in the present study. For example, Paired Observation Number 7 had 0 per cent agreement on patient instigation and independence derived from parallel chunks. This disagreement occurred because one coder decided one "TV" chunk was self-instigated and the other coder decided it was not. The same "TV" chunk caused the disagreement on independence. Similarly, disagreements on patient-instigation and independence derived from parallel chunks in Paired Observation Number 8 were caused entirely by disagreements in the coding of "Watching TV."

Reliability of the Shortform Method

Congruence

Table 7 displays the mean agreement scores for 5-minute segments in the shortform reliability assessment (simultaneous observations using the shortform). Comparison of the paired 5-minute segments within each paired 1-1/2 hour observation resulted in agreement scores ranging from 85.6 to 98.0 per cent and a grand mean agreement score of 93.0 (see Table 7, first column of per cents). The discrepancies within the paired observations again fell into three major categories: (a) differences

Table 7

Agreement Scores (5-Minute Segments):

Shortform Reliability Assessment

Paired Observation Number	Mean Agreement Score (With TV)	Mean Agreement Score (Without TV)
1	98.0	no change
2	93.9	no change
3	96.7	no change
4	85.6	no change
5	91.1	no change
6	96.7	no change
7	91.1	94.4
8	95.6	no change
9	92.8	96.1
10	89.4	93.9
Mean	93.0	94.2

in chunk label based on coder interpretation of observer description; (b) differences in beginning and ending times, although the progression of chunks was similar; and (c) differences attributable to the problem of observing and coding the activity, "Watching TV." All other discrepancies seemed nonsystematic. Removing the effect of the activity, "Watching TV," by considering the other activities the patient was engaged in raised the four affected agreement scores and raised the grand mean agreement score to 94.2 per cent (Table 7, second column of per cents).

Table 8 lists the ten paired observations, ranked in terms of the number of parallel chunks identified within them and the 5-minute segment agreement scores for each paired observation. The Spearman rank order correlation between number of parallel chunks and 5-minute segment agreement score (including the coding of "Watching TV") was -.445, and not significant for N-2=8 degrees of freedom, $\alpha = .01$. Removing the effect of TV chunks slightly strengthened the relationship between number of parallel chunks and agreement scores ($r_s = -.515$, not significant for df=8, α=.01).

Chunks and Accompanying Codes

Table 9 lists the agreement scores on parallel chunks and accompanying chunk codes for the ten paired observations. The first row under the paired observation number lists the number of parallel chunks identified. The remaining rows list the number of parallel chunks for which there was agreement on (a) location; (b) number of different locations; (c) others, if any, who became involved in the activity; (d) instigator(s) of the activity; (e) degree of involvement of the target patient (active, passive, or resistive); (f) others, if any, who conversed with the target patient during the activity; (g) beginning time; and (h) ending time.

The number of parallel chunks in the pairs of observations ranged from 6 to

Table 8

Number of Parallel Chunks and Agreement Scores (5-Minute Segments):

Shortform Reliability Assessment

Paired Observation Number	Number of Parallel Chunks	Mean Agreement Score (With TV)	Mean Agreement Score (Without TV)
1	6	98.0	98.0
8	6	95.6	95.6
3	11	96.7	96.7
7	12	91.1	94.4
9	16	92.8	96.1
5	16	91.1	91.1
6	17	96.7	96.7
10	17	89.4	93.9
4	17	85.6	85.6
2	23	93.9	93.9

Table 9

Agreement Scores on Chunks and Accompanying Codes:

Shortform Reliability Assessment

	Paired Observation Number										Total	
	1	2	3	4	5	6	7	8	9	10	#	%
Number of Parallel Chunks	6	23	11	17	16	17	12	6	16	17	141	
Agreement on (#Chunks):												
Location	6	23	11	17	16	17	12	6	16	17	141	100.0
Number of Locations	6	23	11	17	16	17	12	6	16	17	141	100.0
Others	6	22	10	16	16	17	12	5	15	17	136	96.5
Instigators	6	22	11	17	15	14	10	6	16	16	133	94.3
Degree of Involvement	6	23	11	17	13	17	12	6	16	17	138	97.9
Conversation	5	16	7	11	12	14	9	4	12	14	104	73.8
Beginning Time*	4	21	8	5	15	12	7	6	8	4	90	63.8
Ending Time*	4	21	7	6	13	12	7	6	9	5	91	64.5

*Includes the beginning times of first chunks and ending times of last chunks in each protocol, which are governed by convention (Willems & Crowley, 1976).

23, with a mean of 14 (see Table 9). The total number of parallel chunks for the ten observations was 141. Agreement on location and number of different locations was 100 per cent across pairs of parallel chunks in the comparison. Agreement on the coding of others (if any) who were directly involved in the activity was 100 per cent for five of the ten pairs of protocols, and the mean agreement for this category was 96.5 per cent. Agreement on instigation was 100 per cent for four of the paired protocols, with a mean agreement of 94.3 per cent. Agreement on the patient's degree of involvement was 100 per cent in nine of the paired protocols, and the mean agreement was 97.9 per cent. The coding of persons (if any) who conversed with the target patient yielded a mean agreement score of 73.8 per cent. Lack of agreement in this category was again largely due to the number of instances in which one observer reported conversation and the other did not, rather than instances in which observers disagreed on the identity of a particular person. Finally, agreement on beginning time and ending time of patient activities was 63.8 and 64.5 per cent, respectively.

Fine-Grained Level of Agreement

Table 10 displays the number and per cent of identical line entries in each paired observation (there are ninety lines in each observation). The criterion for agreement between a pair of line entries was that every pair of cells within them be identical. The number of identical line entries in paired observations ranged from 49 to 81 (54.4 per cent to 90 per cent), with a mean of 67.6 (75.1 per cent).

Table 11 presents the number of identical cells within each column (e.g., activity, others, instigation, etc.) in the paired observations. The last row in Table 11 shows the mean percent agreement for each column. This assessment of observer agreement is less stringent than the line entry analysis, but it does pinpoint

Table 10

Number and Per Cent of Identical Line Entries:

Shortform Reliability Assessment

Paired Observation Number	Number of Lines	Per Cent
1	81	90.0
2	50	55.5
3	73	81.1
4	49	54.4
5	57	63.3
6	74	82.2
7	70	77.8
8	80	88.9
9	70	77.8
10	72	80.0
Mean	67.6	75.1

Table 11

Number of Identical Cell Entries in Each Column:

Shortform Reliability Assessment

Paired Observation Number	Columns					
	Activity	Others	Instigation	Involvement	Conversation	Location
1	83	82	83	82	82	90
2	84	81	83	86	56	90
3	89	85	87	89	79	90
4	79	79	80	80	60	88
5	86	87	84	65	82	88
6	88	85	84	88	80	90
7	80	84	80	84	80	90
8	86	86	86	86	82	90
9	84	84	84	82	76	90
10	80	85	85	86	80	90
Total	839	838	836	828	757	896
Mean Percent Agreement	93.2	93.1	92.9	92.0	84.1	99.6

the observer disagreements that reduce line entry agreement. Inspection of Table 11 reveals that on a minute-by-minute basis, observers agreed on the description of patient activity at a level of 93.2 per cent. Similarly, observers agreed on the identity of persons who became involved in patient activity at a rate of 93.1 per cent. Agreement on instigator and degree of involvement was 92.9 and 92.0 per cent, respectively. Conversation, which typically yields the lowest agreement scores in reliability assessments, had a mean agreement of 84.1 per cent. The agreement score for location was extremely high, at 99.6 per cent. Most of the disagreements occurred when the observers were slightly out of phase; i.e., when one observer recorded a piece of information a line earlier or later than the other observer. The second most frequent disagreement occurred when one observer omitted information that the other provided. Only one case of mistaken identity occurred, where observers disagreed on the identity of a person who became involved in the activity of the target patient.

Comparison of Derived Measures

Agreement on derived measures of patient instigation, active involvement, independence, zest, non-idle behavior, two idle behaviors, and one semi-idle behavior was calculated for all of the chunks in the paired observations. Agreement on the same derived measures was also calculated for the pairs of parallel chunks in the paired observations. Table 12 lists the percent agreement scores for all chunks, and Table 13 lists the scores for the parallel chunks. Again, there was very high agreement in the derived measures. The agreement scores for derived measures compiled from parallel chunks were slightly higher than the agreement scores for the derived measures compiled from all chunks, reflecting the deletion of unmatched chunks.

Table 12

Percent Agreement on Derived Measures from All Chunks:

Shortform Reliability Assessment

Percent Agreement on Number of:	Paired Observation Number									
	1	2	3	4	5	6	7	8	9	10
Patient-Instigated Chunks	100	100	100	100	100	100	50	100	100	100
Patient-Active Chunks	100	100	100	100	33	100	0	100	100	100
Independent Chunks	100	100	100	100	100	100	67	100	100	91
Zestful Chunks	100	100	100	100	100	100	0	100	100	100
Non-Idle Chunks	89	100	91	96	100	96	78	100	100	100
"Lying in Bed" Chunks	100	100	93	95	83	100	89	100	91	100
"Sitting" Chunks	100	100	100	100	100	100	100	100	100	100
"Watching TV" Chunks	100	100	100	100	100	100	100	100	100	93

Table 13

Percent Agreement on Derived Measures from Parallel Chunks:

Shortform Reliability Assessment

Percent Agreement on Number of:	Paired Observation Number									
	1	2	3	4	5	6	7	8	9	10
Patient-Instigated Chunks	100	100	100	100	100	0	100	100	100	91
Patient-Active Chunks	100	100	100	100	40	100	100	100	100	100
Independent Chunks	100	100	100	100	100	100	100	100	100	100
Zestful Chunks	100	100	100	100	100	100	100	100	100	100
Non-Idle Chunks	100	100	100	82	100	100	100	100	100	100
"Lying in Bed" Chunks	100	100	100	100	100	100	100	100	100	100
"Sitting" Chunks	100	100	100	100	100	100	100	100	100	100
"Watching TV" Chunks	100	100	100	100	100	100	100	100	100	100

From Table 12 (All Chunks), it can be seen that agreement on patient instigation was 100 per cent for nine of the ten paired observations. Note that the score for Paired Observation Number 6 on patient instigation is 100 per cent on Table 12 (All Chunks) and 0 on Table 13 (Parallel Chunks). Although each observation yielded one patient-instigated chunk, they identified different chunks, which were not paired together.

As in the narrative/shortform comparison, the trend toward high agreement held for the rest of the derived measures in the all-chunk comparison, and agreement usually remained the same or improved when unmatched chunks were deleted. Agreement on active involvement was 100 per cent for eight of the all-chunk comparisons and nine of the parallel chunk comparisons. Extremely low agreement on patient-active chunks occurred in Paired Observation Number 5, in which one observer indicated that the patient was actively involved in one long exercise chunk, while the other observer described the patient as actively involved in a series of five short exercise chunks. Agreement on independent patient chunks was 100 per cent in eight of the all-chunk comparisons and in each of the parallel chunk comparisons. Agreement on patient zest was 100 per cent in nine and ten of the pairings in terms of all and parallel chunks, respectively.

Measures of non-idle behavior, two idle behaviors, and one semi-idle behavior displayed slightly more variability than the measures above. From Table 12 it can be seen that agreement on non-idle chunks ranged from 78 to 100 per cent, and that five pairings yielded agreement scores of 100 per cent in the all-chunk analysis. Deletion of nonparallel chunks (Table 13) yielded nine agreement scores of 100 per cent. In Paired Observation Number 4, one coder combined a series of activities into a single chunk, while the other coder broke the activities into several chunks. The agreement scores for "Lying in Bed" chunks ranged from 83 to 100 per cent in terms of all chunks, with four pairings yielding an agreement score of 100 per cent. Removing nonparallel chunks yielded perfect agreement on "Lying" chunks. There was also perfect agreement on "Sitting" chunks in terms of all and parallel chunks.

"Watching TV" chunks did not seem to present as much a problem as in the narrative/shortform comparison. TV chunks occurred in four of the paired observations (Paired Observation Number 6, 7, 9, and 10), and in terms of all chunks, the agreement on "Watching TV" was 100, 100, 100, and 93 per cent, respectively. Deletion of nonparallel chunks brought all four scores to 100 per cent. Disagreement on the coding of "Watching TV" affected the derived measure, independent chunks, in Paired Observation Number 10 (Table 12). Deletion of nonparallel "Watching TV" chunks improved agreement on independent chunks (Table 13) to 100 per cent.

CHAPTER V

DISCUSSION

When the established narrative system is used as the standard of comparison, the shortform system appears to be both a viable and efficient substitute method. Internally, moreover, the shortform system displays high levels of reliability, or interobserver agreement. Findings relevant to these two issues, as well as some important implications of the study, will be discussed in this section.

The target patient used in the narrative/shortform comparison and the shortform reliability assessment was a quadriplegic who had just begun his rehabilitation program at TIRR. He had begun his exercise programs, but had not begun to sit up or get out of his bed. Many aspects of the findings in this study may be due to the fact that this patient was early in his hospitalization, and was relatively docile; i.e., that the behavior streams of early patients typically produce fewer chunks than the behavior streams of patients ready for discharge from the hospital. The problem of low frequency was frequently encountered in this study, and was more apparent in the shortform reliability assessment than in the narrative/shortform comparisons.

Narrative System versus Shortform System

The shortform data are highly congruent with the data from the narrative system. In the narrative system, reliability scores of 90 per cent (5-minute segments) are considered satisfactory. The grand mean agreement score (87.8) for 5-minute segments of narrative/shortform paired observations was acceptable, particularly considering that this comparison involved data from different observers <u>and</u> different systems. Discrepancies between narrative and shortform data due to differences in chunk labels were not damaging in this comparison because discrepancies of this nature are inherent in the narrative system itself. Discrepancies due to the approximated beginning and ending times of chunks in the shortform system were expected, and will be discussed below. Differences attributable to the problem of observing and coding the activity "Watching TV" became so obvious that agreement scores were calculated on data both with and without the "Watching TV" chunks. The problem of handling data on TV and suggested solutions will be discussed below. None of the other differences appeared to be systematic, and none could be attributed to any difference in the systems because the disagreement rates did not exceed typical intra-narrative system reliability rates.

Agreement on most of the accompanying codes for parallel chunks was extremely high. Codes for location and number of locations agreed 100 per cent; however, these codes seldom prove troublesome in the narrative system. Agreement on others involved in the patient activity, instigation, and the patient's degree of involvement were also quite high. None of the differences in these codes were attributable to differences between the narrative and shortform systems.

The coding of conversation seemed to present difficulties--the agreement score was only 67.6 per cent. It should be noted that the test for agreement on conversation was quite stringent. The codes for conversation in parallel chunks were judged to agree only if all of the components of these codes were identical. For example, in a set of parallel chunks, if one observer listed conversation with a visitor, medical technician, and physician, while the other observer listed conversation with a visitor and medical technician, the codes for conversation were judged to disagree completely. In most cases, disagreements in conversation codes occurred when one observer reported conversation and the other did not, rather than instances in which observers disagreed on the identity of a particular person. None of the disagreements could be attributed to coder omission of information that the observer had provided. The problem with coding conversation seemed to be primarily at the data-collection or observer level of both the narrative and shortform systems. This code is the weakest code in both systems (see Dreher, 1975, for typical agreement scores in the narrative system), and presented problems for both narrative and shortform systems.

Although the agreement on beginning time and ending time was extremely low, these codes did not appear especially troublesome. Agreement on these codes reflected the application of the coding conventions used to define the beginning time of first chunks and the ending time of last chunks (Willems & Crowley, 1976), and the rare occasions in which patient behavior began or ended on the minute or half-minute. Disagreements, however, could be attributed to differences in (a) time as reported by observers and (b) the coding system. Differences in time as reported by observers occurred frequently in the narrative system, and seemed to reflect differences in observers' styles.

Three classes of disagreement on time were directly attributable to differences in the coding systems: (a) time convention; (b) chunk on form, S on narrative; and (c) gap. The other classes of disagreement were routine problems in the narrative system itself and were not strictly attributable to differences between the two systems. These routine problems comprised almost two-thirds of the seconds of disagreement.

Discrepancies due to time conventions accounted for almost one-third of the seconds of disagreement. However, the seconds of disagreement due to time convention comprised only 2.1 per cent of the total seconds of data collected (1108 out of 54000). These disagreements occurred with greatest frequency at the junction between parallel chunks and reflected the degree to which the paired observations were out of step. Taken in the context of the great congruence between the two systems and the high agreement on chunks and accompanying codes, this source of error seems to be one of minimal importance. In short, the error due to difference in time conventions appears relatively insignificant.

The second class of disagreement has been called, "Chunk on form, S on narrative." Almost one-sixth of the seconds of disagreement occurred where a chunk in form data was paralleled by an S or series of Ss in the narrative data. This sort of error was of particular interest because it affected agreement on number of chunks and on some of the derived measures. In the narrative system, S-codes are termed "marginal" codes because they are attached (in the margin of the typed protocol) to chunks; the inclusion or deletion of S-codes does not affect the underlying chunk itself. When a particular fleeting bit of patient care was given an S in the narrative and omitted in the shortform, the two systems still agreed if the underlying chunks agreed. If, however, the bit of patient care was given an S in the narrative and was chunked in the shortform, the two systems disagreed--even if the other underlying chunks agreed. In other words, agreement existed only if both systems coded the patient care as a chunk or if neither coded it as a chunk.

The shortform system did not have the option of assigning an S or series of Ss to highly intermittent and sporadic patient-care activities. One reason that some activities were combined into chunks in the shortform and not combined in the narrative is that coders may have felt compelled to account for the patient care activities, and the only code possible was a chunk code. A second reason may be that shortform observers themselves combined the bits of patient care into a single summary entry. In any case, this sort of discrepancy accounted for a large proportion of the disagreement between the two systems, and is an important source of error.

The third class of discrepancy, labeled "Gap," proved to be an insignificant source of disagreement between the two systems. Even though gaps in coding are permitted in the narrative system and not in the shortform, this type of discrepancy accounted for only .5 per cent of the total disagreement. Further, gaps usually occurred at the junction between parallel chunks, and did not affect the congruence between the two systems or the agreement on chunks. Beginning time and ending time are the only chunk codes affected by this class of discrepancy; the remaining chunk codes and the derived measures produced by them would be unaffected.

Agreement scores on derived measures in general were quite high. It should be noted, however, that (a) many of these measures are interrelated and (b) that the extremely low frequencies in some measures spuriously inflated or deflated some agreement scores. For example, consider the derived measure, Patient-Instigated Chunks. This measure is related to the number of non-idle chunks, because only non-idle chunks receive codes for instigator(s). Idle chunks receive the code "Irrelevant" for instigator(s). Generally, coders have greater opportunity to code the patient (or anyone) as instigators of activities as the number of non-idle chunks in a patient record increases. Further, certain kinds of behavior are easily broken down into several episodes rather than tied together in a single chunk. "Watching TV," "Reading," and "Transporting" are prime examples of stop-and-start behaviors that may be coded as one or several chunks. When a coder elects to break an activity into several discrete chunks, he must then select and code an instigator for every new chunk. "Watching TV" was extremely troublesome in this analysis. Disagreements on instigator(s) were generally confined to cases in which one coder elected to use one patient-instigated TV chunk, while the other coder used three or four patient-instigated TV chunks. Because "Watching TV" affected this and other measures to such a great extent, it will be discussed separately below.

Another problem with the analysis of patient-instigated chunks was the extremely low frequency of such chunks. The target patient was early in his comprehensive rehabilitation program at the hospital, and early patients are typically inactive and docile. Five of the ten pairs of patient records agreed completely that there were no patient-instigated chunks. All but two of the disagreements in the remaining pairs of observations involved the coding of "Watching TV" as one long or several short chunks. Those two disagreements involved the question of whether or not the patient verbally instigated some of the things done with him by other persons. A coding convention in the narrative system was also used in the shortform system; i.e., coders selected the patient as instigator only if it was clear that he had asked that something be done immediately before it began (Willems & Crowley, 1976). The two disagreements here involved two single instances in which someone scratched the target patient, and in both cases the chunks dealing with this activity did not have parallel chunks in the other system.

The derived measure, "Patient-Active Chunks" is also related to the number of non-idle chunks. Only non-idle chunks may be coded for the patient's degree of involvement (active, passive, or resistive). The coding of patient involvement is a highly inferential task, especially in the narrative system. Coders must imagine a continuum of possible involvement levels for any particular activity, and then locate the patient's degree of involvement on that continuum. In the shortform system, observers must perform that inferential task and provide an involvement code on a minute-by-minute basis. Coders then select and apply the modal score to the chunk. Although coding degree of involvement is separated from observation in the narrative system and is an integral part of the observation in the shortform system, agreement on this measure was quite high. The extremely low frequency of active involvement by this target patient must also be considered. Seven of the observation pairings agreed completely that there were no patient-active chunks. In three pairings, one record yielded one more active chunk than the other. Overall, however, agreement on this measure was quite high.

The measure, "Independent Chunks" is the number of chunks which are instigated by the patient and in which he is unassisted. This measure is closely related to the number of non-idle chunks and the number of patient instigated chunks. Seven of the ten pairings agreed that there were no independent chunks. The other three pairings disagreed on the number of independent chunks that occurred, and in every case, the chunks in question were "Watching TV." The problem occurred when one record used several short TV chunks and the other used a single long TV chunk. In this analysis, it became apparent that coding TV-watching activities presented special problems for the derived measure of patient independence. These problems and a proposed solution will be presented below.

The measure, "Zestful Chunks," or patient zest, is the number of chunks which are instigated by the patient and in which he is actively involved. This measure then, is closely related to the number of non-idle chunks, number of patient-instigated chunks, and number of patient-active chunks. Agreement on this measure was extremely high; nine of the ten pairings agreed that there were no chunks displaying patient zest. In the last pairing, one observer reported subtle but active patient involvement in an arm exercise while the other reported that the patient was not working at all. This points up the inferential problem of coding the level of participation in exercises by a severely handicapped person. There is very little functional musculature in a quadriplegic's arms, and many normal cues (such as flexing biceps) are not present. Observers must occasionally use intuition to decide if a patient is expending great energy, and coders must rely on the observers for good, in-context information. This is a normal problem in the analysis of patient behavior, but it should serve as a reminder to both observers and coders that they must be extremely vigilant in reporting and coding patient behavior.

Measures of the number of non-idle, idle, and semi-idle behaviors displayed more variability than the measures above. Much of the variability in the per cent agreement on number of non-idle chunks was due to the relatively small number of such chunks. Further, deletion of nonparallel or unmatched chunks greatly improved agreement between the two systems. Most of the extraneous or unmatched chunks were short episodes of "Watching TV," or cases in which an activity received an S or series of Ss in the narrative and a chunk code on the shortform. The number of non-idle chunks is a critical measure, however, in calculating certain ratios of active, self-instigated, independent, or zestful patient behavior. For this reason, it is extremely important to continue to assess the similarities and differences in these measures derived from the narrative and shortform systems, by observing a target patient much closer to discharge, and therefore more active, than was the target patient used in the present study.

With an early patient, the number of non-idle chunks is usually similar to the number of idle chunks; i.e., non-idle chunks are usually singular occurrences between two idle chunks. The coding of idle chunks, aside from beginning time and ending time, is governed completely by convention. There are few decision points for observers and coders with early patients because the proportion of idle to non-idle behaviors is high. An early patient often produces long idle chunks interspersed with single non-idle chunks, but an advanced patient often produces many non-idle chunks in a row. Findings from other studies indicate that a patient closer to discharge would be involved in more non-idle activities, and would provide more decision points for both observers and coders (LeCompte & Willems, 1970; Vineberg & Willems, 1971; Willems, 1972a, 1972b; Willems, Crowley & Dreher, 1974; Willems & Vineberg, 1969a, 1969b).

Overall, agreement on the idle and semi-idle behaviors was high. Generally, high agreement on non-idle chunks was matched with high agreement on "Lying in Bed" chunks. One reason for this is that the target patient was early in his career at TIRR, and his activities usually alternated between single idle and non-idle episodes. Another reason is that the target patient was confined to his bed, so that "Lying in Bed" was the best label available to coders when the patient was idle. All ten paired observations agreed that there were no "Sitting" chunks. Strictly speaking, "Watching TV" is not an idle behavior because it does require watching a TV screen and listening to an audio speaker. However, such marginal behavioral involvement is difficult to observe and code because of its passive and docile nature. Seven of the paired observations agreed that no TV watching occurred. However, whenever "Watching TV" behavior occurred, there was never perfect agreement on the number of "Watching TV" chunks. The problems involving TV activities will be discussed below.

Reliability of the Shortform System

The data obtained from the shortform system proved to be highly reliable across all pairs of observers, and the grand mean agreement score on 5-minute segments (93.0) compared favorably to agreement scores yielded in assessments of reliability in the narrative system (Dreher, 1975). Discrepancies between observation pairs due to (a) differences in chunk label and (b) differences in beginning and ending time (although the progression of chunks was similar) seemed to be routine problems in the shortform system. These discrepancies arose because observers and coders were required to summarize the patient's activities and to approximate beginning time and ending time. Despite these types of disagreement, the shortform seemed to have good reliability for many kinds of patient activities. "Watching TV," which caused difficulty in the narrative/shortform comparisons, did not present an obvious problem in the shortform reliability assessment.

Agreement on most of the accompanying codes for parallel chunks was again extremely high. Codes for location and number of locations agreed 100 per cent. Agreement on others involved in the patient activity, instigation, and the patient's degree of involvement was also quite high. Every difference in these codes was attributable to differences in the reports made by observers. Conversation again proved to have the lowest agreement score. Lack of agreement was largely due to the number of instances in which one observer reported conversation and the other did not; none of the disagreements were coders' errors of omissions. This code remained one of the weakest codes in the narrative and shortform systems.

Agreement on beginning time and ending time was also very low. There are several possible reasons for the lack of agreement on this measure. First, observers were not given explicit instructions about time conventions. One observer reported that she had decided to record certain activities in the half-minute in which they began (e.g., 30 seconds of "Brushing Teeth" that began at 1:25:45 and ended at 1:26:15 was reported for the second half of the minute 1:25), while another observer reported that she assigned those activities to the half-minute in which they ended. When conventions covering instances like this are finalized, this sort of discrepancy should be eliminated. Second, observers were using watches that could not be synchronized. The inability to set two watches to the same second is not a problem in the narrative system, because observers report running clock time and the constant difference between the two records can be added to or subtracted from every time entry in one of the observations after the records have been transcribed. In the shortform system, however, paired observations must begin at precisely the same time, because each minute's summary of patient behavior must be derived from exactly the events.

In an attempt to synchronize their observations, pairs of observers met before the observation to determine the difference in the readings on the watches they used. The "slow" watch became "standard time," and each minute began for the observer with the slow watch when his watch read 0 seconds; e.g., 1:12:00. If the difference between the watches was 15 seconds, the observer with the "fast" watch had to remember that standard time began for him when his watch read "15 seconds after," or 1:12:15. This caused a great deal of confusion for the observer who had to make the time correction, especially during periods of great activity, and no doubt contributed to the lack of agreement on beginning time and ending time.

Third, when the ending times of parallel chunks disagreed, the beginning times of the next chunks disagreed also. No gaps were permitted in the shortform system. Nonetheless, agreement scores on beginning and ending times in the shortform system compared very well with reliability assessments on these measures in the narrative system.

The high agreement scores for 5-minute segments, parallel chunks, and accompanying codes were reflected by high agreement in the fine-grained analysis of the paired shortform observations. Over 75 per cent of the line entries (seven cells in each line) were in complete agreement. Further, agreement within each column (activity, others, instigation, degree of involvement, conversation, and location) was also extremely high. This analysis helped to pinpoint the source of differences in the assessment of 5-minute segments, parallel chunks, and accompanying codes. Most of the disagreements occurred when the observers were slightly out of phase; i.e., when one observer's record listed a piece of information one line earlier or later than the other observer's record. A few disagreements occurred when one observer omitted information that the other observer provided. Only one case of mistaken identity occurred, where observers disagreed on the identity of a person who became involved in the activity of the target patient. None of the disagreements was attributable to coder error--in every case, coders provided correct code numbers for the information recorded by observers.

As in other comparisons, agreement on the derived measures was generally quite high. As mentioned above, however, many of these measures are interrelated, and extremely low frequencies spuriously affected some agreement scores. In general, there were more chunks per observation in this analysis than in the narrative/shortform analysis, but the relatively low frequencies underlying the important derived measures makes this analysis rather weak.

The measure, "Patient-Instigated Chunks" produced very high agreement in both overall and parallel chunk comparisons. Six of the paired observations agreed that there were no patient-instigated chunks. In two of the paired observations, disagreement on the number of patient-instigated chunks was confined to the issue of coding "Watching TV" as one long versus several short chunks. The remaining paired observations contained disagreement on the coding of patient care activities which were requested (verbally) by the target patient. Clearly, the coding of "Watching TV" and activities that are verbally instigated by the target patient are persistent problems.

The measure, "Patient-Active Chunks," although highly inferential at the observational level, produced high agreement. Eight of the paired observations yielded no patient-active chunks. In one of the remaining observations, one coder broke an exercising episode into five segments, while the other coder left it as one. In every exercising chunk the patient was coded as active. Although agreement on this measure appears dismally low (33 per cent), the observers themselves agreed that the patient was actively involved in most of the exercising episode. Therefore, low agreement in this case was not directly attributable to differences in the information provided by the observers, but to differences in the decisions made by the coders.

The measure, "Independent Chunks" produced extremely high agreement. Seven of the paired observations agreed that there were no independent chunks. One pair of observations yielded an identical number of independent chunks; all of these chunks were "Watching TV." Two of the paired observations yielded disagreement on the number of independent chunks and in every case, "Watching TV" was the chunk involved. Again, the problem occurred when one protocol listed several short TV chunks while the other had a single long TV chunk.

The measure, "Zestful Chunks," or patient zest, had very high agreement scores. Nine of the ten pairings agreed that there were no chunks displaying patient zest. The last pairing involved an activity which was coded as a half-minute chunk in one observation and not reported at all by the other observer. If the assumption that both observers were vigilant is valid, then the observers may have been using different rules-of-thumb in making the decision to report or delete what must have been a rather brief activity. If this is the case, then formal guidelines for observers will help eliminate disagreements of this nature.

The measures of non-idle chunks displayed high agreement. Deletion of nonparallel or unmatched chunks (mostly "Watching TV") improved agreement on number of non-idle chunks to 100 per cent in nine of the paired observations. The remaining disagreement involved the exercising episode broken into five chunks in one observation and left intact in the other. Nonetheless, this agreement rate is remarkably high. Similarly, agreement on "Lying in Bed" chunks, "Sitting," and "Watching TV" chunks was very high. Deletion of nonparallel and unmatched chunks brought agreement rates up to 100 per cent in every case.

In summary, the reliability of the shortform system seems exceptionally strong, even when assessed by measures originally developed for the narrative system. At almost every conceivable level--cell-by-cell, minute-by-minute, chunk-by-chunk--the data display very high agreement. The rigor of the reliability assessments and the results obtained in the evaluation of the shortform compare very favorably with the results and accepted standards in other observational research (Bijou, Peterson, & Ault, 1968; Hawkins & Dotson, 1975; Johnson & Bolstad, 1973; McDowell, 1973; Weick, 1968; Wright, 1960, 1967). It may be that the measures used here are not stringent enough to pinpoint single cases of serious disagreement. However, if used routinely in an ongoing shortform system, these measures might help to reveal (a) troublesome activities that require special observer attention; (b) troublesome activities that require special coding conventions; (c) systematic drift over time in observers' notations; and (d) systematic drift over time in coding.

Savings

The shortform system was very efficient and allowed much quicker turn-around time. Most of the shortform records were final coded within 1-1/2 hours of completion of the observation. The late-night observations were usually coded the next day. Although coders were never reminded of their work or urged to hurry and complete it, all ten records were ready for transfer to keypunch sheets two days after the last observation. In contrast, few of the narrative records were coded until three or more days after the observation was completed. All of the narrative records were ready for transfer to keypunch sheets within a week of the last observation. The time required for transfer to keypunch and computer analysis is two to four days (these tasks are performed on a contract basis by another office). The turn-around time required to carry one week's work from the first observation through computer analysis would average 10 days for the shortform system and 14 days for the narrative system. In other words, the data presented at patient rounds (cumulative, quantitative descriptions of patient performance--pace, mobility, independence, zest, rate of active behaviors, etc.) would be 10 to 3 days old in the shortform system and 14 to 7 days old in the narrative system. The data yielded by the shortform could provide more immediate feedback to the treatment staff at patient rounds.

The shortform system allowed immense savings in person hours. While observers themselves invested the same amount of time in observation and working with coders, typing and rough coding time were eliminated completely. Final coders reported that they invested approximately the same amount of time in coding, but that they spent much less time tracking down people to get answers they needed to complete the coding. Checkers required less time to complete their tasks, and the time needed to transfer the data to keypunch sheets was the same as that in the narrative system. In the narrative system, approximately six different persons must become involved in the (a) observation, (b) typing, (c) rough coding, (d) final coding, (e) checking, and (f) transferring to computer keypunch sheets. After the initial observation, about 12 person hours are required to process the data. In contrast, the shortform requires, at most, four persons to complete the (a) observation, (b) final coding, (c) checking, and (d) transferring to keypunch sheets. After the initial observation, only two person hours are required to process the data in the shortform system.

The shortform system also allowed savings in office supplies and equipment. The shortform does not require tape recorders, cassette tapes, typewriters, typing ribbon, tape transcribers and all of the associated paraphernalia. The ten narrative observations, transcribed, used 121 pages; the ten shortform observations required 60 pages. In sum, the shortform offers savings in supplies, equipment, and person hours while providing increased systems efficiency.

Overall Assessment

The shortform system seems to be a viable alternative to the narrative system of patient observation. The shortform system displayed a great deal of congruence with the narrative system in terms of 5-minute segments, parallel chunks and their accompanying codes, and derived measures. The shortform also proved to be very reliable when tested for agreement on 5-minute segments, parallel chunks and their accompanying codes, minute-by-minute data entries, and derived measures. The shortform also afforded considerable savings in person hours and other resources. Suggested guidelines for observers and coding conventions are presented in Appendix C.

One major difference between the narrative and shortform systems became apparent, and wholesale adoption of the shortform system should be delayed until the ramifications of this difference can be fully explored. One of the great strengths of the narrative system is its separation of data gathering (observation) and data management (coding). This separation allows the system to be heuristic. That is, coding conventions; categories of persons, places, and activities; and derived measures may be added to or deleted from the gathering of raw data when necessary. Observers provide raw material only; refinement of the data is completed in coding and computer analysis. Because observers provide only raw material in the narrative system, all of the decisions regarding the identification of chunks, others involved in chunks, instigators, locations, etc., are made by coders. Further, this separation of observation and coding requires the participation of many different persons to carry an observation through transfer to keypunch--and each new person checks the work of all those who precede him.

In the shortform system, however, observers provide less raw data for the coders to work with. Observers decide for themselves what the principal activities of the target patient are, and then provide all of the necessary accompanying information for that activity. Despite the blurred distinction between data gathering and data reduction and management, the two systems appear highly congruent, based on ten paired observations. However, the congruence between these systems may change over time, because the narrative strictly separates observation from coding while the shortform does not. Further evaluations should be completed to assess change in congruence between the two systems over time.

Recommendations

Four recommendations regarding the use of narrative and shortform systems in behavioral research were developed as a result of the present study. The first deals with the problems associated with the observation and coding of "Watching TV," and is based on the findings of the present study. The second recommendation deals with the problem of assessing intercoder agreement, when the target patient is mostly idle, or docile. The last two recommendations deal with the application of the narrative and shortform system in other settings, and to digital data acquisition systems.

Observation and Coding of "Watching TV"

The analyses of 5-minute segments, chunks, and derived measures presented in the preceding chapter strongly suggest that the activity of watching TV presents problems for observers and coders in both systems. First, it is usually very difficult for observers to discriminate between times when the patient is watching TV and other times when the patient is idle in bed, particularly when the patient is in head tongs and traction. One problem is that the observer must determine whether or not the patient is actually looking at the TV screen and listening to the audio speaker. If the observer is able to provide this information each time the patient looks at the TV (and often he is not), the coder must decide how much TV watching behavior is necessary to code the activity as a chunk. The extremely sporadic nature of TV-watching behavior makes this a difficult task.

Furthermore, "Watching TV" chunks affect some derived measures in an undesirable way. The measure, "Independent Chunks" (self-instigated and performed alone) was intended to be a measure of patient activity. TV-watching behavior (which is usually self-instigated and performed alone) more closely resembles idleness than activity. The present coding of TV-watching behavior is not congruent with the basically idle nature of this activity.

Perhaps in the future, observers should continue to report TV watching as accurately and completely as possible. The coding of this activity should be changed in both the narrative and shortform systems so that the patient's behavior stream is accurately represented and the integrity of the derived measures is preserved. Rather than breaking TV-watching behavior into various chunks that interrupt or overlap the patient's other activities, coders could chunk whatever else the patient is doing and attach a marginal code to denote the presence of concurrent TV-watching behavior.

Re-Examination of the Reliability Assessments

One by-product of this study has been a new look at some of the reliability assessments used in the narrative system. One measure of reliability that has been used frequently is the number of seconds of agreement expressed as a percentage of the entire 90-minute observation. A much stronger assessment of reliability would take into account only non-idle chunks or chunk minutes. Rates of non-idle behaviors represent the crucial measures of behavior and feedback to the hospital staff. It is suggested, therefore, that the following measure be added to the battery of reliability assessments:

$$X = \frac{2\text{ (number of agreed-upon non-idle seconds)}}{\text{(number of non-idle seconds in one protocol + number of non-idle seconds in other protocol)}}$$

Application of the Narrative and Shortform Systems to Other Settings

Both observers and coders reported that the transition from narrative system to shortform system was not difficult, although most reported some initial awkwardness. At first, observers were reluctant to summarize patient behavior, but after a little practice, most felt comfortable with the new mode. Observers and coders all agreed that their prior experience as coders helped them more than their experience as observers (most were cross-trained in the narrative system). In view of the combined observation-coding required of observers in the shortform system, coder training in the narrative system will undoubtedly be a necessary part of observer training in the shortform system.

At this point, it does not seem appropriate to pick up the shortform system and apply it too hastily to all problems that require behavioral observation. Even though it functions with high reliability, the shortform is derived wholly from a specific narrative system, and has been evaluated with respect to that system alone. If a shortform system is desired, the following groundwork should be completed first, and then the shortform can be used with some confidence.

1. Delineation of the specific problem area, target population, and behavioral focus. Formulation of research resources. Training of personnel.
2. Short-term narrative observation of target populations and behaviors. Derivation of coding categories, conventions, and possible behavioral measures. Continued training of personnel.
3. Short-term narrative observation to test the adequacy of coding categories, conventions and derived measures. Development of reliability assessment techniques. Continued training of personnel.
4. Overlapping narrative and prototypical shortform systems. Comparison of the systems. Modifications made where necessary. Continued training of

personnel.

5. Full-scale use of shortform and narrative systems in tandem. Continued reliability, validity assessment.
6. Phasing into the shortform system.

Application of the Narrative and Shortform Systems to Digital Data Acquisition Systems

Electronic digital systems for recording behavior in many settings are now available from commercial sources. Behavioral data can be punched on keyboards and stored on magnetic tape, which in turn can be played into a computer, keypunch, or digital printer. Code frequencies, sequences, real-time durations, and modified frequency and time-sample scores can be extracted from these records by hand or by computer. In many ways, the use of digital data acquisition systems would offer great flexibility to many programs of observational research.

Use of a digital data acquisition system in the narrative system would be difficult, but not impossible. The first stage amenable to computerization is final coding--codes for chunks could be entered at a terminal instead of the margins of typed pages of narrative description. This would, however, make checking a more arduous and confusing task, since checkers must read the text and check every code against the text. Filling out transfer-to-keypunch sheets is a chore that would also be amenable to digital acquisition systems, and would eliminate the need to pay for keypunch services. This possibility should be explored.

Adaptation of the shortform system to digital data acquisition systems would be relatively simple, and probably very costly. The following system would, however, greatly reduce turn-around time. Observers could use a hand-held keyboard to enter alpha-numeric data on a minute-by-minute basis, just as they do in the pencil-and-paper shortform. These data could be stored on magnetic tape, and could be displayed on a terminal screen for observer and coder to review and edit. The alpha-numeric codes provided by the observer could be translated by either the coder or by a computer program. The coder could designate the beginning time, ending time, and label for each chunk in a single-file sequence, and a computer program could derive the rest of the accompanying codes from the data in the cells between beginning and ending times. If chunks overlapped, the coder could elect to identify all the chunks and accompanying codes himself. Next, the checker could review both the observation record and final coding at the terminal screen. If no questions arose, the checker could send the observation record to the printer, and the final coding first to keypunch and next to the printer. The simplicity of the shortform system makes it suitable for this kind of processing, which would reduce turn-around time to a matter of hours.

BIBLIOGRAPHY

Altmann, J. Observational study of behavior: Sampling methods. Behavior: An International Journal of Comparative Ethology, 1973, 44, Parts 3-4, 227-276.

Barker, R. G. Observation of behavior: Ecological approaches. Journal of Mt. Sinai Hospital, 1964, 31, 268-284.

Barker, R. G. Explorations in ecological psychology. American Psychologist, 1965, 20, 1-14.

Barker, R. G. Ecological psychology. Stanford, Calif.: Stanford University Press, 1968.

Bijou, S. W., Peterson, R. R., & Ault, M. H. A method to integrate descriptive and experimental field studies at the level of empirical concepts. Journal of Applied Behavior Analysis, 1968, 1, 175-191.

Dreher, G. F. Reliability assessment in narrative observations of human behavior. Unpublished masters thesis, University of Houston, 1975.

Hawkins, R. P., & Dotson, V. A. Reliability scores that delude: An Alice in Wonderland trip through the misleading characteristics of interobserver agreement scores in interval recording. In E. Ramp & G. Semb (Eds.), Behavior Analysis: Areas of research and application. Englewood Cliffs, N. J.: Prentice-Hall, 1975. Pp. 359-376.

Johnson, S. M., & Bolstad, O. D. Methodological issues in naturalistic observation: Some problems and solutions for field research. In L. A. Hammerlynck, L. C. Handy, & E. J. Mash (Eds.), Behavior change: Methodology, concepts, and practice. Champaign, Ill: Research Press, 1973. Pp. 7-67.

LeCompte, W. F., & Willems, E. P. Ecological analysis of a hospital: Location dependencies in the behavior of staff and patients. In J. Archea & C. Eastman (Eds.), EDRA-2: Proceedings of the 2nd annual environmental design research association conference. Pittsburgh: Carnegie-Mellon University, 1970. Pp. 236-245.

McDowell, E. E. Comparison of time-sampling and continuous-recording techniques for observing developmental changes in caretaker and infant behaviors. Journal of Genetic Psychology, 1973, 123, 99-105.

Vineberg, S. E., & Willems, E. P. Observation and analysis of patient behavior in the rehabilitation hospital. Archives of Physical Medicine and Rehabilitation, 1971, 52 (1), 8-14.

Weick, K. E. Systematic observational methods. In G. Lindzey & E. Aronson, (Eds.), Handbook of social psychology (2nd Ed.). Vol. 2. Reading, Mass.: Addison-Wesley, 1968. Pp. 357-451.

Willems, E. P. The interface of the hospital environment and patient behavior. Archives of Physical Medicine and Rehabilitation, 1972, 53, 115-122. (a)

Willems, E. P. Place and motivation: Complexity and independence in patient behavior. In W. J. Mitchell (Ed.), Environmental design: Research and practice. Los Angeles: University of California at Los Angeles, 1972. Pp. 4-3-1 to 4-3-8. (b)

Willems, E. P. Behavioral ecology and experimental analysis: Courtship is not enough. In J. R. Nesselroade & H. W. Reese (Eds.), Life-span developmental psychology: Methodological issues. New York: Academic Press, 1973. Pp. 195-217.

Willems, E. P. Longitudinal analysis of patient behavior. In W. A. Spencer (Ed.), Annual report of research on Research and Training Center No. 4. Houston, Texas: Texas Institute for Rehabilitation and Research, 1975. Pp. A-57-A-112.

Willems, E. P. Behavioral ecology, health status, and health care: Applications to the rehabilitation setting. In I. Altman & J. Wohlwill (Eds.), Environment and behavior: Advances in research. Vol. 1. New York: Plenum, in press.

Willems, E. P., & Campbell, D. E. Behavioral ecology: A new approach to health status and health care. In B. Honikman (Ed.), Responding to social change. Stroudsburg, Pa.: Dowden, Hutchinson and Ross, 1975. Pp. 200-210.

Willems, E. P., & Crowley. L. R. Narrative recording, structural coding, and content analysis of human behavior. Houston, Texas: Texas Institute for Rehabilitation and Research, 1976.

Willems, E. P., Crowley, L. R., & Dreher, G. Direct observation of patients in the rehabilitation hospital: Assessment of change after three years. Final report of Project No. R-135. In W. A. Spencer (Ed.), Annual Report on Research and Training Center No. 4. Houston, Texas: Texas Institute for Rehabilitation and Research and Baylor College of Medicine, 1974. Pp. A-93 to A-123.

Willems, E. P., & Raush, H. L. (Eds.) Naturalistic viewpoints in psychological research. New York: Holt, Rinehart and Winston, 1969.

Willems, E. P., & Vineberg, S. E. Direct observations of adults with spinal cord injuries. (12 Vols.) Houston, Texas: Texas Institute for Rehabilitation and Research, 1969. (a)

Willems, E. P., & Vineberg, S. E. Direct observation of patients: The interface of environment and behavior. Psychological Aspects of Disability, 1969, 16, 74-88. (b)

Willems, E. P., & Vineberg, S. E. Procedural supports for the direct observation of behavior in natural settings. Houston, Texas: Texas Institute for Rehabilitation and Research, 1970.

Wright, H. F. Observational child study. In P. H. Mussen (Ed.), Handbook of research methods in childhood development. New York: Wiley, 1960. Pp. 71-139.

Wright, H. F. Recording and analyzing child behavior. New York: Harper and Row, 1967.

APPENDIX A

Coding Manual for Narrative System

CODING SYSTEM: DEFINITIONS, GUIDELINES, AND CONVENTIONS

Page

CODING SYSTEM: DEFINITIONS, GUIDELINES, AND CONVENTIONS

I. FOCUS

Assume that the protocol captures and describes, in a continuous manner, the ongoing experience-behavior stream of a patient, in which the major parts can be (a) something the patient does (in the active sense); (b) something that is done to, happens to, or is done with the patient; and (c) segments in which the patient is overtly idle or passive. The protocols are not well suited to detailed documentation of what others (staff, other patients, visitors) do, per se. Thus, focus is continually upon occurrences in the experience-behavior stream of the target patient. The presence of others and what they do will only serve the purpose of adding refinements and supporting codes to the analysis of patient behavior.

II. CHUNKS

A. Given these assumptions, mark off the protocols in terms of behavior stream events, or chunks. A chunk is a molar occurrence in the behavior stream of the target patient that meets the following criteria.

1. A chunk can be readily characterized (and, therefore, labeled) by a single principal activity, e.g., Taking Medication, Using Bedpan, Being Bathed, Standing on Standing Board, Playing Chess, Sitting in Wheelchair, Exercising Arms, Eating Lunch, Lying in Bed.

2. A chunk's principal activity (a) begins at a clearly described starting point, (b) occurs over time (sometimes short) in a characteristic, regular manner, with essential supporting and accompanying activities (see Convention V-C) until (c) a clearly described stopping point, which may be signaled by a goal achieved, a task finished, a tailing off of the principal activity, or an interruption.

3. One absolutely necessary condition for marking a chunk is the direct involvement of the target patient. It must be one or more of the following: (a) something the target patient is doing; (b) something being done to him; or (c) something being done with him. On this criterion, no matter how detailed and flowery the observer's description of it, a checker game between two OTs (while the target patient is engrossed in working leather off to the side) would not be marked as a chunk, whereas Working on Leather would be. In other words, the focus, the key to looking for and marking chunks is always the behavior stream of the target patient.

4. Verbal behavior, or what is often called conversing, is never to be marked as a chunk no matter how deeply or extensively the target patient is involved. (See IV-A below for coding of verbal interchanges.) Rather, during the time that the patient is involved in verbal interchanges, be sure to chunk what he is doing by way of nonverbal behavior, e.g., Sitting in Wheelchair, Standing on Standing Board, Transporting to OT, Playing Cards.

5. Things the patient is actually doing, whether idle or not, present little problem. Things that are done to him or with him should be coded as chunks only if they directly and immediately affect his body through bodily contact, the laying on of hands, or moving or jostling him, his clothing, or equipment on him (be sure to see Convention V-B below).

6. Several chunks can occur simultaneously, or overlap each other. For example, some of the following may overlap in various ways or occur simultaneously: Having Pulse Taken plus Having Temperature Taken, Standing on Standing Board plus Reading, Being Irrigated plus Watching TV, Drinking Coke plus Studying Chart.

7. Assuming regular, continued occurrence of the principal activity throughout the period of the chunk does not necessarily mean that one specific behavior is continuing without interruption throughout. A principal activity, and therefore, a chunk, may have fluctuations, variations, and complexities. For example, the principal activity, Exercising Arms, can and probably will include resting in between, exchanging information, verbal feedback, modifying equipment, essential instructions, and perhaps other things (See V-C below).

B. The best procedure for chunking is to browse through an extended piece of the protocol (perhaps 3 or 4 pages at a time) and ask, in broad strokes, "What is happening here by way of principal activities?" or "How would I characterize this section in terms of molar, principal activities?" It is usually quite helpful if, during the browsing, you make short-hand notes, circle or underline names of persons and events, and make marks that will serve as cues in the process of chunking and rating. After this global browsing and marking, proceed to the actual marking of chunks (see Figure I).

C. Next, mark the actual chunks that occur in the section. Your orientation toward chunking should be to mark off major, molar events in the patient's behavior and then to ask, "Why should I divide the sequence up any finer?", rather than to mark off tiny, molecular events and then ask, "Why should I combine any of these into larger units?" The former orientation will help to avoid the proliferation of tiny events. After you are confident that the chunks are marked appropriately, proceed by attaching all accompanying codes to them. NOTE: Even though molar units of behavior often take longer than molecular ones, molarity is not defined by lapsed time. Some molar events can be quite short; e.g., Combing Hair, Transferring from Mat to Wheelchair (especially by a late patient). Thus, molar units are defined by the trajectory, integrity, and continuity of principal activities and not by time.

D. A chunk is marked on the protocol by short horizontal lines extending into the lefthand margin. The two horizontal lines that mark the beginning and end of a chunk should then be connected by a continuous vertical line to form a major bracket to which chunk labels and refining codes will be attached. Whenever a chunk spans the change from one page to another, put an arrow on the vertical line at the bottom of one page and at the top of the next. Whenever a chunk continues beyond a page, write a part of the chunk label at the top of the subsequent page. When two chunks occur simultaneously at any point, the chunk that begins first is the primary chunk; the chunk that begins later is the overlapping chunk. If a chunk overlaps another, designate it in the margin as Overlap-1. If a chunk overlaps two others at any point, designate it as Overlap-2 (see Figure 2).

5:37:00	The patient continues to eat his lunch. He is alone; no one is near him. He is in Station III.
5:37:30	The patient continues to eat.
5:37:42	The patient puts down his fork and pushes the dinner tray away.
5:37:51	The patient sits quietly now.
5:37:55	The patient reaches over to his bed and picks up the book there. He converses with NA James as he does this.
5:38:12	The patient opens the book. Conversation ends. The patient is turning pages in the book.
5:38:25	The patient begins to read.
5:38:50	NA James walks up beside the patient. The patient continues to read.
5:39:10	NA James puts a thermometer in the patient's mouth. Reading continues.
5:39:30	The patient continues to read. The thermometer is in his mouth. He converses briefly with James.
5:40:00	No change.
5:40:35	He stops reading and looks around the room.
5:40:45	NA James removes the thermometer.

Figure 1

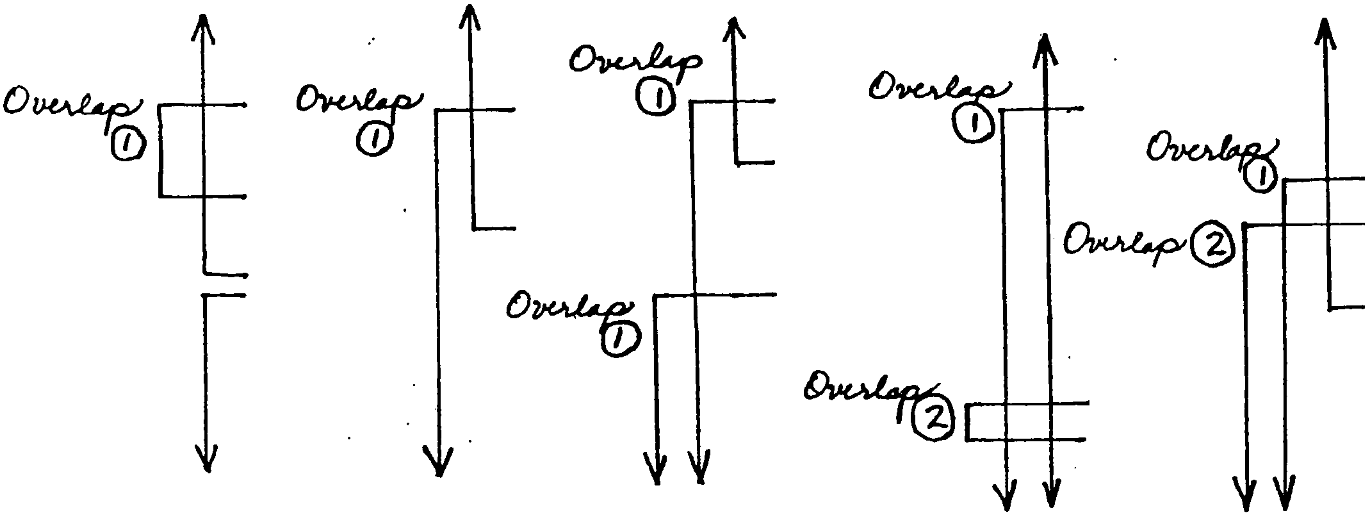
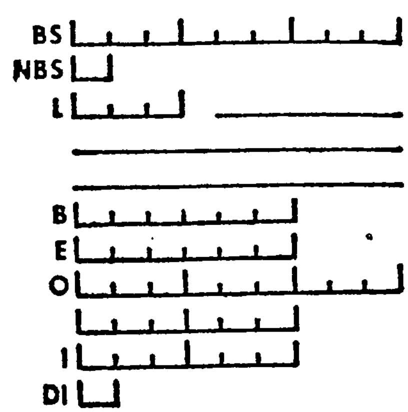


Figure 2

III. CHUNK CODES

Next to the bracket for every chunk, write in the following items of refining and clarifying information. The format for this information is provided by a rubber stamp with blanks to be filled in as follows:



A. Behavior Settings (BS). Enter three-digit code number(s) indicating the behavior setting(s) in which the chunk takes place. If more than one setting is involved, list them sequentially, even if the same setting is involved twice. For example, if a patient (during one chunk) goes from Station 4 to Lobby and back to Station 4, you should enter the three-digit codes for Station 4, Lobby, and Station 4 in sequence (Station 4 twice). If more than three are involved, list (a) the setting of origin (where the chunk started), (b) a principal middle setting (the major transitional setting is usually Hallways), and (c) the setting of termination (where the chunk ended). If many settings are involved in one particular chunk, it may help if you make a reference list in the margin of the protocol. Be sure to acquaint yourself with the TIRR Behavior Setting Survey, because the inclusion of areas within settings often does not correspond to architectural boundaries and sometimes violates common sense and intuition.

B. Number of Behavior Settings (NBS). Use a one-digit number to indicate the number of different settings involved in the chunk. For example, in the above case (Station 4 to Lobby to Station 4), the number of different settings is 2. "9" is used to indicate any amount over eight settings.

C. Chunk Label (L). Write a label describing the principal activity of the chunk. Use your own wording, but make the label as informative as possible, though short. Do not provide the chunk classification codes; these will be filled in independently. If the chunk is judged to have begun before the observation period started, enter a "1" in the first space of the L code. You may use the introductory material only to determine whether or not a chunk is ongoing at the beginning of the protocol. If the chunk is judged to continue past the end of the protocol, enter a "2" in the first space of the L code. If a chunk is judged to start precisely at the beginning of the protocol, or to terminate precisely at the end of the protocol, enter a "0" in the first space of the L code.

D. Beginning Time (B). Write in the chunk's beginning clock time, using six digits: two for hour (on the 24-hour, or military clock), two for minute, and two for second.

E. Ending Time (E). Write in the chunk's ending time in a manner similar to beginning time.

F. Other Persons Directly Involved in the Principal Activity of the Chunk (O). The criterion for noting someone here is that he or she must be clearly described as being directly and behaviorally involved in the principal activity of the particular chunk in question, e.g., therapist conducting or helping with treatment. The code should be made in terms of categories of persons, using the prepared list, Categories of Persons. For each category of persons directly involved, use a three-digit number. If more than one category of others is involved, enter the three-digit numbers in the appropriate places.

1. In coding involvement by others the requirement is that a person be an essential component of the principal activity of the chunk. You should ask, "Would the execution of the principal activity be affected if the person in question were not there?" If the answer is yes, code the person O; if the answer is no, do not code the person O. By this rule, the person who puts a thermometer in a patient's mouth is an O in Having Temperature Taken. A person who brings in a meal tray and sets it down is not coded O for Eating unless he or she becomes directly involved in Eating by helping with it.

2. The criterion for coding involvement by others is not temporal in nature; i.e., the requirement is not that a person be directly involved in the principal activity throughout the entire period of the chunk. Rather, you should ask, "Who, if anyone, was directly involved in the principal activity?", and code the categories of persons who were; e.g., the type of person involved in Having Temperature Taken, even though that person participates only by putting in and taking out the thermometer. By this criterion, a person who becomes directly involved in Eating by giving a patient only one of many bites of food is coded O.

3. 300 (Cannot be Inferred or Irrelevant) is applied whenever (a) someone else is directly involved, but his classification cannot be inferred from the record or (b) the principal activity is some form of behavioral idleness.

4. 400 is used for chunks in which more than five (5) classifications of persons clearly participate in the activity.

5. 290 (No others involved) is applied whenever a patient engages in a nonidle chunk without the direct involvement of anyone else. NOTE: Never use 010 (patient himself) for an O code.

G. Who Instigated or Initiated the Principal Activity of the Chunk (I). This refers to the person at whose initiative, instigation, or arrangement the principal activity started or got rolling. This code will be indicated by a three-digit number from the prepared list, Categories of Persons. If two categories of persons instigated the chunk, enter both.

1. In coding instigation, ask yourself, "Who, if anyone, started this activity?" You may use both behavioral and verbal cues to identify the instigator(s), but always look for behaviors or verbalizations that indicate the most immediate instigation of the chunk in question. The clearest cues are behavioral in nature; use verbal cues only when you can ascertain that (a) they did, in fact, cause the chunk to occur, and (b) they immediately precede the chunk.

2. 300 (Cannot be Inferred or Irrelevant) is always applied to (a) chunks whose principal activities and labels center around behavioral idleness, e.g., Lying in Bed, Sitting in Wheelchair, Waiting; and (b) chunks in which the classification of the instigator(s) cannot be inferred from the record (see conventions V-J and V-L for coding I for Standing on Standing Board, Progressive Sitting, and Watching TV).

3. 400 is used for chunks in which more than two classifications of persons clearly share in the instigation. If none of the instigators is the target patient, use 400 alone; if the target patient himself is also an instigator, use 400 and 010.

4. 010 (patient himself) should be used for instigation when the patient instigates the activity. NOTE: Do not use 290 (No other persons involved) for the I Code.

H. Degree of Involvement in the Principal Activity of the Chunk by the Target Patient (DI). This refers to the general level of the patient's participation in the principal activity. Be sure to examine the chunk in its entirety for references to the patient's level of participation. Degree of involvement may vary throughout the chunk, and you must decide which of the following four codes is best suited to the patient's general degree of involvement.

Level 1. Active involvement. Apparently helping along or participating actively in the principal activity. When rating degree of involvement, it is important to envision a variable scale that characterizes the maximal and minimal degree of involvement or effort possible for different activities. The patient's degree of involvement will be coded at Level 1 when he participates with apparent enthusiasm, gusto, or high expenditure of energy, according to the nature of the activity in question. Thus, for example, even though the patient's expenditure of energy might be lower while Playing Dominoes (moving pieces, etc.) than while submitting passively to Having Arms Exercised by a PT, he would be rated DI-1 for Playing Dominoes and DI-2 for Having Arms Exercised. The question of rating Level 1 versus Level 2 arises most often when someone else is directly involved in the principal activity and arises less often when the patient is doing something alone. In any case, give a chunk Level 1 when you have some clear, positive reason from the record to code it this way.

Level 2. Passive involvement. Apparently only cooperating in going along with the principal activity; it is being done to him and he is essentially passive; he is doing it with little apparent enthusiasm or gusto. Often, when the patient is doing or responding to something that someone else has physically arranged or started, you will give it Level 2. NOTE: Whenever you cannot infer the degree of involvement and there is no indication of resistance by the patient, give the chunk Level 2.

Level 3. Resistance. Apparently offering some difficulty, active resistance, or strenuous complaint against the activity, even if he finally does part or all of it.

Level 4. Degree of involvement, as a code, is irrelevant. NOTE: Always apply Level 4 to chunks that involve some form of complete behavioral idleness as their principal activities.

In choosing a code for degree of involvement, apply the one that characterizes the patient's degree of involvement generally during the chunk. In a case that is truly equivocal between Levels 1 and 2, apply Level 1. Use Level 3 only when the patient's participation is largely and consistently resistive.

SUMMARY: Be sure to use the zero digit in filling in all chunk codes (e.g., 101 for setting, 09-02-15 for time, 300 for other, 010 for instigation). You should enter material into eight places on the stamped format, leaving blank only the three-digit number code for classification of the chunk label, and any unused 3-digit places for setting, O, or I codes.

IV. MARGINAL CODES

Two kinds of occurrences, which will not be part of the chunk codes, will be noted in the margin of the protocol when they occur.

A. Mutual Verbal Interchanges (C). If it is clear from the protocol that the target patient is involved in a mutual verbal interchange, indicate the occurrence in the margin.

1. Verbal interchanges are never coded as chunks.

2. The criterion for indicating these events is that the patient actually be involved in a mutual exchange, with some back-and-forth, some give-and-take of verbalization. Such mutual exchanges may be very short.

3. When a mutual verbal interchange occurs, indicate it in the margin with a capital C and a three-digit number for the kind of other person involved.

4. Code C with each distinct type of other person only once per chunk. Thus, if during a given chunk (e.g., Eating Lunch), the patient is involved in mutual verbal

interchanges with three LVNs at one time or with three different LVNs at three different times, you should code only one C-026.

5. Within the span of a given chunk, enter a new C code and its three-digit number for each new type of other person with whom the patient engages in a mutual verbal interchange.

6. Begin anew with the coding of Cs for each new chunk, even when a mutual interchange is obviously continuing from a previous behavioral chunk.

7. Draw a line from each C code to the vertical bracket line of each chunk to which the C is attached. (a) If a C occurs while two chunks are overlapping, draw lines from the C to both chunks. (b) If a C occurs during a gap between chunks, draw the line to the chunk that best locates the C by place and time.

B. Fleeting Bits (S). If someone makes a short, fleeting incursion into the patient's behavior stream, code it in the margin, subject to the following criteria.

1. Even though the event is very short, it must meet the same criteria of directness that apply to chunks in which someone does something to or for the patient (See Convention V-B below). Many times, chunks which do not meet the 15-second criterion (Convention V-E below) will be coded as S-bits.

2. It must be (a) a behavioral, nonverbal bit in which the person does something that is clearly and positively directed to the patient's care, comfort, welfare, treatment, protection, or well-being (e.g., fluffs up patient's pillow and leaves, unbuttons patient's shirt that is too tight); or (b) a behavioral, nonverbal activity to which the patient responds in some behavioral manner. S-bits must be clearly directed to the patient; physical contact alone is not sufficient reason to code an S. If the activity in question is neither clearly and positively directed to the patient's care, comfort, welfare, treatment, protection, or well-being, nor clearly mutual, do not code an S. Examples of events that may be coded as S: shaking hands with the patient; giving the patient pills; feeling patient's head for fever. Examples of events that may not be coded S: putting reciprocals in patient's lap; patting patient on the arm; holding hands with patient.

3. Never code essential or necessary parts of chunks as bits. If you are considering coding a bit during a chunk, ask yourself whether the principal activity of the chunk would suffer, or be detracted from, if the bit had not occurred. If the answer is yes, you should probably not code the behavior in question as a separate bit, but consider it part of the ongoing chunk.

4. Behavior of the target patient is never coded as a bit.

5. Indicate a bit by writing an S in the margin, together with a three-digit number indicating the type of other person involved in it; e.g., S - 111. Draw a line from each S code to the vertical line of the chunk during which the S occurs. (a) Never connect an S to more than one chunk. If an S occurs while two chunks are overlapping, draw one line to the chunk that best locates the S by time and place.
(b) If an S occurs during a gap between chunks, draw the line to the chunk that best locates the S by place and time.

6. Every S must be coded. If more than one type of person is involved, enter an S and the appropriate 3 digit number for each type. You may also record the same type of person more than once per chunk, by writing an S and the appropriate code number for every occurrence.

C. Remember that it is easy to overlook C and S codes because they are not included in the rubber stamp format for chunks. Be vigilant in spotting and coding Cs and Ss.

V. SOME COMMENTS, RULES-OF-THUMB, AND CONVENTIONS

A. Labels for Chunks. The labels for chunks should be phrased from the perspective of the target patient at all times, with the "ing" verb form wherever possible; e.g., Being Bathed, Exercising Arms, Talking on Telephone, Lying in Bed, Sitting in Wheelchair, Having Temperature Taken, Eating Lunch in Cafeteria. Attach your own descriptive label and do not use category labels from the chunk classification system.

B. Chunking What is Done To or For the Patient. The problem here is that you will often seem to be chunking what someone other than the patient is doing. When chunking something that is being done to or for the patient, demarcate as chunks only those behaviors by others that immediately and directly affect or implicate the patient through (a) direct physical contact by touching his body or the clothing or equipment that is on him (not just connected to him); or (b) moving him or jostling him directly (not indirectly, as in simply changing the orientation of the surface he is on); or (c) direct behavioral involvement with the patient when the patient participates in the principal activity. Some problems here will be dealt with later.

1. The major exceptions to condition (b) will be transports. Transfers will usually not be exceptions because the patient will be lifted, pushed, or pulled directly. (See V-I for conventions regarding transports and transfers.) Another exception will be turning and repositioning a patient in a Stokes-Manville bed. This bed turns mechanically from side to side, so Being Turned or Being Repositioned chunks will many times include episodes in which the patient is not moved or jostled directly.

2. Examples of events that will usually be chunked by these criteria are:

Having intravenous apparatus attached (condition a)
Being helped with a drink (condition c)
Being bathed (conditions a and b)
Being transferred (conditions a and b)
Having bed clothes changed while in bed (conditions a and b)
Having blood pressure checked (condition a)
Being turned (conditions a and b)
Having leg bag emptied or checked (condition a)

3. Examples of events that will usually not be chunked by these criteria are:

Having bedside table cleaned up
Having bed bag checked or emptied (the urine bag hanging on the bed frame)

Having someone check the IV bottle hanging on the stand, or change its rate of flow
Having someone bring a meal tray and put it by the bed
Having bed lowered or raised (this will not be chunked when it occurs by itself. It might often be included in the activity involved in transfers.)
Having bed pushed closer to wall (this will not be chunked when it occurs by itself. It might often be included in the activity involved in transfers or transports.)

C. An Activity as a Reasonable or Essential Part of a Chunk. What might often appear as an activity distinct enough to be a separate chunk might not be. If a "new" activity is an essential part, an intrinsic part, of the principal activity of a chunk already marked, it should not be marked as a separate chunk, e.g.:

"Aide James is bathing patient Joske. James turns Joske onto his right side. James starts to bathe Joske again." The turn is an essential part of the chunk, Being Bathed, in that James must do this to continue and finish the bath. Therefore, it is not marked separately from Being Bathed. Other examples: shifting and moving patients during the course of a treatment, or tightening equipment during an exercise.

A principal, molar activity (and, therefore a chunk) often includes a number of stops, starts, apparent changes in behavioral pace and orientation, and apparent changes of and sequential reciprocity of involvement. Within the general criterion of molar behavioral trajectory and integrity of a principal activity, the inclusion of "reasonable, essential, supportive accompaniments" must often be construed quite broadly. In other words, when coding molar behavior, one should not infer personal, psychological goals and intentions. At the same time, however, concepts such as the following can be helpful: (a) behavioral consistency, integrity, or trajectory toward some observable goal or end point; (b) normal behavior perspective (reasonable inference regarding what is involved in a behavioral act for a given person); (c) apparent goal directedness across the topography of an act, with reasonable fluctuations within the act. This should clear up many questions regarding when to mark extended activities as one chunk or more than one chunk when chunks subsume a number of ebbs, flows, and behavioral variations. For example, Exercising Arms may include many essential, supportive activities and rest periods that are part of it by this convention; Having a Shampoo may include washing, drying and combing hair.

1. Ambiguous Movements and Changes of Location Between Exercises. Separate chunks will be marked when the modality of exercises changes (see V-C-5). However, patients sometimes move around or are moved around between two types of exercise. Code such movements as separate chunks if (a) they involve changing location (across mats, to different spot on mats) and not simply turning over (turning from stomach to back or reversing position of head and feet) and (b) they occur between two distinctly different chunks. When an event meets these criteria but does not exceed 14 seconds (see Convention V-E), always include it in the subsequent chunk. If the event does not exceed 14 seconds and the subsequent chunk is an idle chunk, drop the event in question completely.

2. Ambiguous Preparatory Activities. Never label a chunk "Preparing for ________." Rather, label all chunks in terms of their principal activities. In some cases (e.g., between transporting and an exercise, or between transferring and an exercise), an

extended patient activity might occur which, in some persons' eyes, could be seen as preparation for a subsequent chunk; e.g., Putting on Pads and Braces (between Transporting to PT and Walking on Parallel Bars), Putting on Equipment (between Transferring to Mats and Exercising Arms on Pulleys). When such activities occur, proceed as follows:

(a) Think in terms of principal activities and determine whether the event meets the criteria of a chunk or whether it should be included in the previous chunk or subsequent chunk.

(b) If the event does seem to stand separately, then chunk it separately if 1) the target patient does it alone <u>or</u> 2) all the other persons directly involved in it are different from the other persons directly involved in the subsequent chunk, <u>or</u> 3) it is separated from the chunk for which it is ostensibly preparatory by another, interspersed chunk (e.g., Lying on Mats, Sitting in Wheelchair). In other words, chunk it separately if it fulfills <u>any</u> of these three conditions.

(c) Even if the event seems to stand separately, always include it with the subsequent chunk for which it is ostensibly preparatory if 1) someone else is directly involved <u>and</u> 2) at least some of the other persons involved are the same in the two chunks <u>and</u> 3) it is not separated from the chunk for which it is ostensibly preparatory by another, interspersed chunk. In other words, include it in the subsequent chunk only if it fulfills <u>all</u> of these three conditions.

(d) See Convention I-1 for coding activities that are ostensibly preparatory for transfers and transports.

3. <u>Ambiguous Tail-End Activities</u>. At times, some activities at the transition points of chunks will be ambiguous. For example, when someone unbuckles straps, fusses with clothing, and jostles and rearranges the patient's equipment at the end of a standing board activity, it is sometimes difficult to decide whether these activities should be included in the standing board chunk or the transfer chunk; i.e., finishing Standing on Standing Board vs. Transferring (in the sense of getting ready for transfer).

(a) When such activities can, in fact, be seen as the tail-end of one chunk, include them in that chunk if they directly follow (are attached to) that chunk before a change to another chunk. For example:

1) Standing on Standing Board--(the activities in question)--Lying on Standing Board

2) Standing on Standing Board--(the activities in question)--Being Transferred to Wheelchair

In both cases, include the activities in the standing board chunk.

(b) When such activities are separated from the previous chunk by a new chunk, then include them in the later transfer chunk. For example:

> Standing on Standing Board--Lying on Standing Board--(the activities in question)--Transferring to Wheelchair

In this case, include the activities in the transfer chunk.

4. Ambiguous Movements at the End of an Exercise Period. Patients will often move or be moved to the edge of the mats in PT at the end of an exercise period, as if preparing to transfer. Include such movements with the subsequent transfer chunk unless they are separated from it by another chunk; i.e., Lying on Mats. If the event does not meet the 15-second rule (Convention V-E) in the latter case, code it as an S unless the patient moves by himself.

5. Chunks in Treatment and Nursing Care. Occasionally you will have to decide whether to code a sequence in a treatment facility or on the ward as one or more chunks. When the treatment modality changes (arm exercises with wrist cuffs to arm exercises with elbow extension devices; arm exercises to leg exercises), code separate chunks. When the nursing modality changes (temperature to pulse; changing dressing to medication), code separate chunks. Important clues in both cases are changes in (a) bodily modality, (b) props and equipment, and (c) instructions, rules of the game, and sequencing.

6. Troublesome Grooming Activities and Personal Hygiene. Often, a patient (especially one who does much for himself) will string several specific acts together, e.g., washing face, brushing teeth, combing, in such a way that it would make sense to code the whole sequence together as one Grooming or Personal Hygiene chunk. Follow these guidelines: (a) If the specific acts are actually clumped together in place and time, make them one chunk, with one label. (b) If the specific acts are separated either by gross place or by other, clearly different, activities, code separate chunks.

7. Behaviors as Chunks vs. Intermittent Pieces. A few types of behavior can either occur with enough concerted continuity and integrity to be coded as chunks or occur as intermittent pieces widely dispersed over time and across other behaviors. For example, Drinking a Soft Drink can occur in a relatively concerted, continuous fashion so that there will be no doubt about chunking it (even allowing for reasonable pauses and accompanying movements). On the other hand, occasional brief sips from a soft drink may occur over an extended period of time across another activity or activities (e.g., Playing Pool, Standing on Standing Board, Working on Leather, or even Sitting in Wheelchair and Looking Out of Window). In the latter case, chunking the drinking will involve some strain. Other examples of behavior about which such a problem can arise are drinking coffee, eating peanuts, or eating potato chips. The problem with these behaviors is that, even though they can often occur as chunks, their pieces can sometimes occur intermittently during the principal activities of other chunks. (See Convention L for coding Smoking.)

(a) If such a behavior meets the criteria of a chunk, then code it as a chunk. Allow for reasonable and necessary accompaniments, such as pauses and appropriate movements. If it does not meet the criteria of a chunk, and no one else is involved in it, drop it.

(b) Here are some concepts or principles that might help in making the decision:

1) Continuity, consistency, and integrity toward some end point in the ordinary behavior perspective. Just as chewing and swallowing one pickle is not Eating Lunch, so taking one sip, or consuming two potato chips (when done in isolation, during another chunk) are not Drinking Coke, or Eating Chips.

2) Molar perspective. Ask what the patient is doing in the molar sense in the protocol section under question. If he is basically Transferring, Playing Pool, Sitting in Wheelchair, or Lying in Bed, and only occasionally (incidentally) taking a sip or consuming a mouthful, then chunk what he is doing in the molar sense and drop the fleeting, intermittent consummatory events.

3) Single-object necessity and continuity. When you chunk drinking a Coke (or coffee), or eating chips it should be an event in which the patient, of necessity, consumes all or a major part of that specific item (i.e., that Coke) until some stopping point. Even though it appears somewhat awkward to argue in this fashion, disconnected, highly intermittent sips, and mouthfuls could, in principle, each involve different cups, and different chip bags that were somehow placed around the environment. Whenever the separate pieces of such a series could, in principle, involve new items, do not mark the series as a chunk.

(c) When such an event is chunked, someone else might often be coded as an O; e.g., someone else holding the cup for the drink, or giving the chips or peanuts to the patient to eat. When an intermittent series of unchunked pieces of behavior occurs, someone else might still be directly involved in each. In that case, attach marginal S codes.

D. Time Judgments. There are several sticky problems involved in judging beginning and ending time for chunks. Figures 3 through 9 illustrate the issues and conventions discussed in the text which follows.

1. When a chunk is judged to end at (just before) a time entry (e.g., 4:21:15), be sure to give the chunk full time, which would be up to the crucial time entry (4:21:15). This will anchor many ending times of chunks. (See Figure 3)

2. The bracketing of chunks will usually produce back-to-back chunks (chunk A will end at the point where chunk B will begin). At these times, the ending time for A will be the same as the beginning time for B. (See Figure 4)

3. Chunks sometimes begin or end during the descriptive material for a single time entry, and therefore, there is no clear time available to anchor either the beginning or ending. When only one end of a chunk is not anchored by a time entry, assume you are randomizing error by using the midpoint between the time entry just before and the time entry just after the critical point as your anchor for time. Calculate the midpoint by the following formula:

$$\text{Midpoint} = \text{beginning time} + \frac{\text{end} - \text{begin}}{2}$$

rounded to the closest even second. (See Figure 5)

4. When a chunk is judged to begin some time after one entry and end some time before the next, use the midpoint between the two time entries to establish beginning time and use the second time entry to establish ending time. (See Figure 6)

5. When you judge a chunk to be occurring, but there is no clear description of its end, establish ending time by using the beginning time of the next chunk if the next chunk is nonidle. (See Figure 7) If the next chunk is idle, continue the ambiguous nonidle chunk for one minute past the last timed reference to the chunk and begin the idle chunk at that point. (See Figure 8) When the chunk in question (e.g., Drinking Coke, Reading Magazine) is overlapping within another, much longer chunk (e.g., Standing on Standing Board), establish ending time for the chunk in question by ending it one minute after the last protocol reference to it. (See Figure 9)

6. Always bracket chunks first and leave the bracketing intact; i.e., do not allow the distortions that result from these time conventions to change your bracketing.

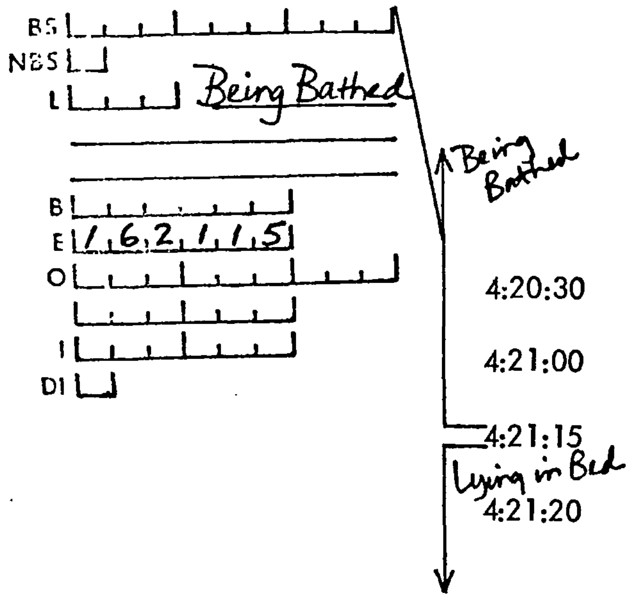
E. Arbitrary Minimum Length of Chunks (the 15-second rule). An arbitrary minimum length of 15 seconds will be adopted to facilitate coding and avoid the accumulation of tiny chunks. Thus, whenever the subtraction of beginning time from ending time for an event is 14 seconds or smaller, do not chunk the event, no matter how dramatic it is.

1. Sometimes, when this rule eliminates an event, it will be meaningful to combine behaviors in new ways into chunks. For example, imagine a patient in Station 4 engaging in the following sequence of activities: (a) Transporting to Frig; (b) Getting Orange Juice from Frig; and (c) Transporting to Room. If (a) and (c) are 12 seconds each, you might possibly code the whole sequence as one chunk (Getting Orange Juice from Frig) and include the movements in it.

2. Often, no new meaningful combination of behaviors into chunks will be possible. In such a case, do not strain. Rather, simply leave unchunked and blank the span of material for the event that is too short. (see Convention V-F.)

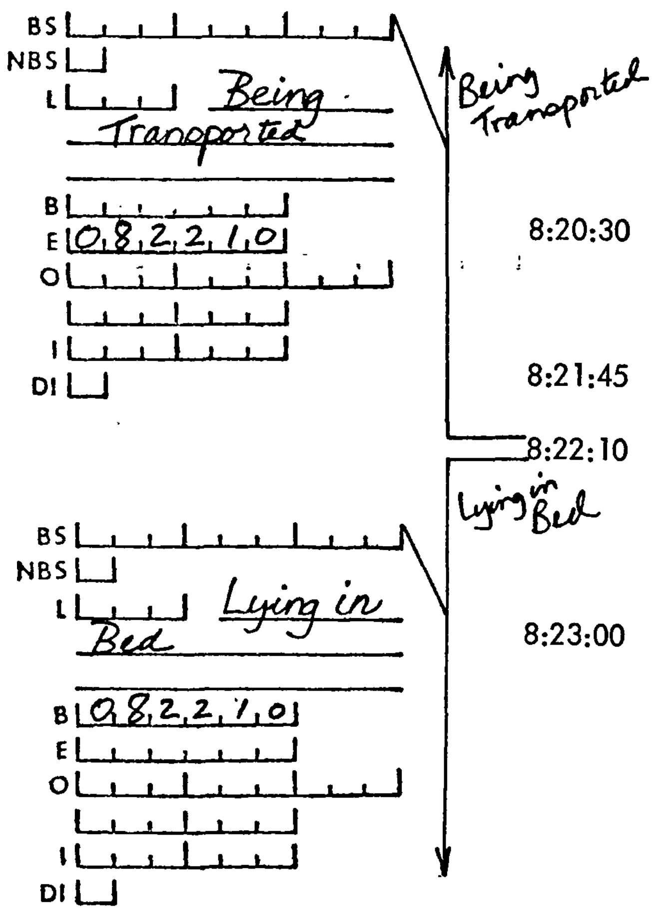
3. If an event is shorter than 15 seconds and meets the criteria of an S-bit, it may be coded as an S.

4. This 15-second rule supercedes all other conventions.



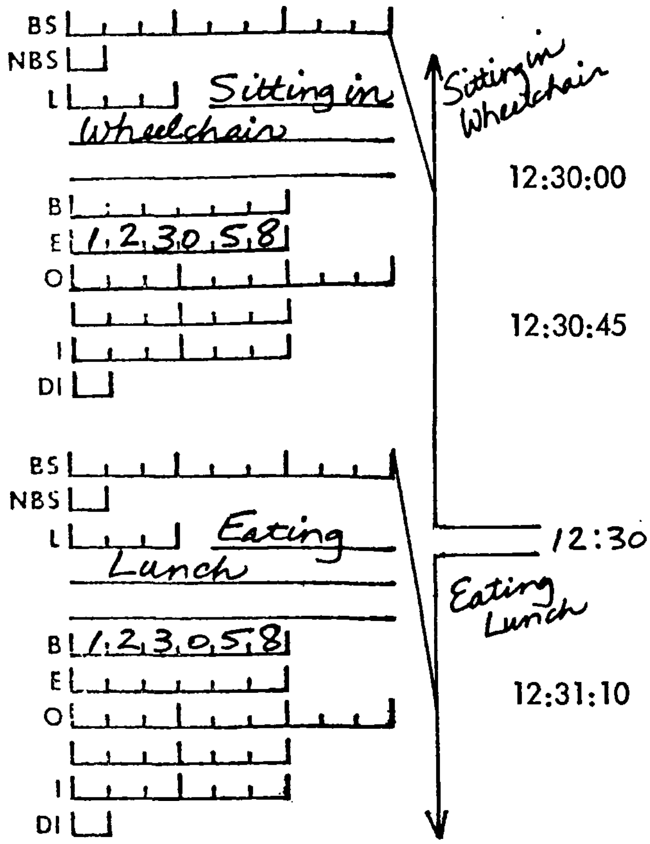
4:20:30	Patient is being bathed by the nurse.
4:21:00	Bathing continues.
4:21:15	The nurse stops bathing the patient.
4:21:20	The patient lies quietly in his bed.

Figure 3



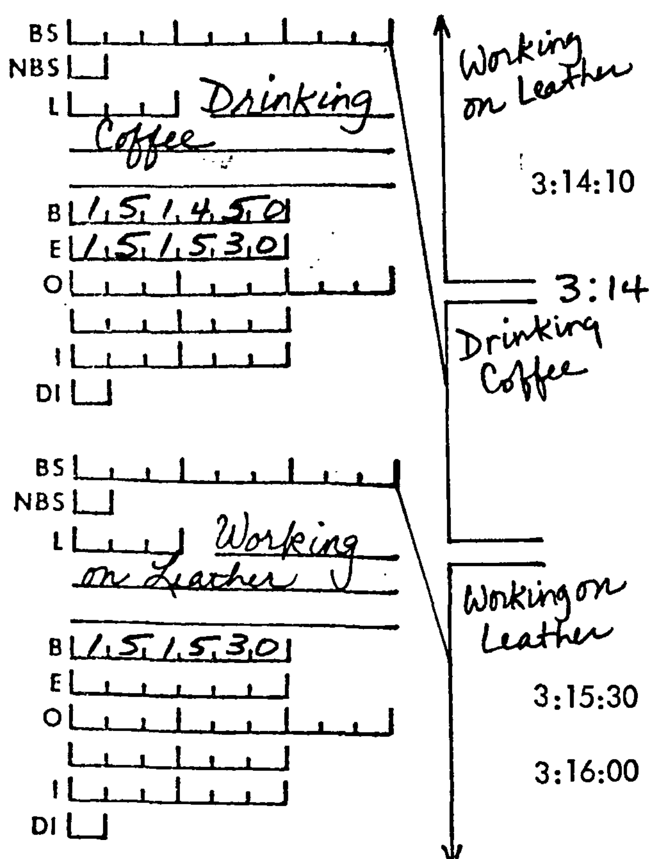
8:20:30	The transport continues, as the patient is being pushed down the hall in his bed.
8:21:45	Transport continues.
8:22:10	The transport ends. The orderly leaves the bedside. The patient is lying quietly in his bed.
8:23:00	The patient closes his eyes.

Figure 4



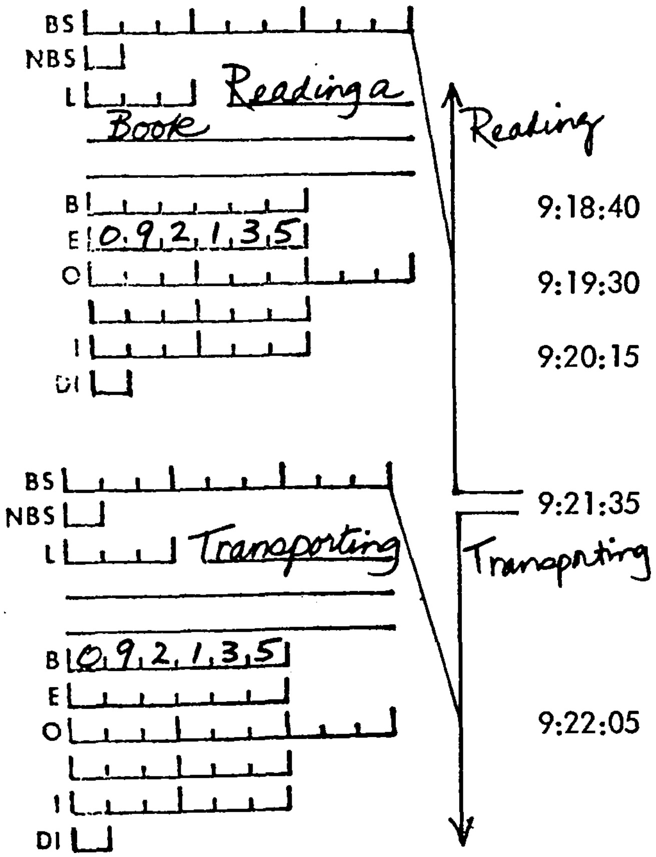
12:30:00	The patient is sitting quietly in his wheelchair.
12:30:45	The patient is sitting quietly in his wheelchair. The nurse brings his lunch tray. She opens it.
12:30:58	She begins to feed the patient. She gives him his salad first.
12:31:10	The patient continues to eat his lunch.

Figure 5



3:14:10	The patient is working on leather. He puts the awl down and picks up his coffee cup. He is slowly drinking the coffee. He pauses, then takes another long, slow drink. He has finished his coffee. He picks up the awl and begins to work on the leather again.
3:15:30	The patient is working on leather.
3:16:00	No change.

Figure 6



9:18:40 The patient is reading a book.

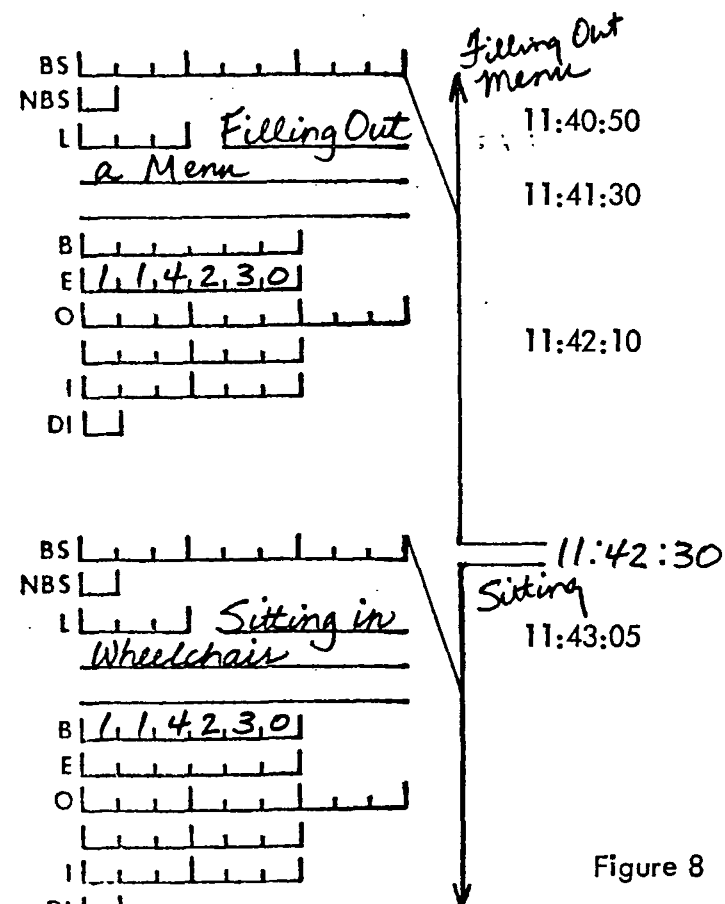
9:19:30 The patient continues to read.

9:20:15 The patient turns a page and continues to read.

9:21:35 The patient begins to wheel his wheelchair out of his room and into the hall.

9:22:05 The patient continues to wheel down the hall.

Figure 7



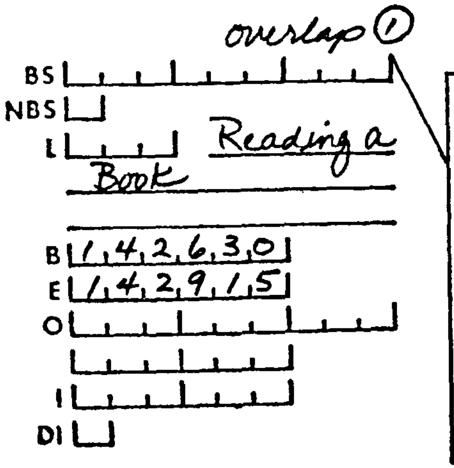
11:40:50 The patient is filling out a menu.

11:41:30 The patient continues to read the menu and to write on it.

11:42:10 The patient is sitting quietly in his wheelchair, looking at the ceiling. He now looks out the door.

11:43:05 The patient continues to sit in his wheelchair.

Figure 8



Standing On Standing Board

2:25:45	The patient continues to stand on the standing board. He has a book on the work table on the board.
2:26:30	The patient begins to read the book.
Reading	
2:27:45	The patient turns a page in the book.
2:28:15	The patient continues to read. He turns another page.
2:28:45	The patient continues to stand on the standing board.
2:29:15	
2:29:20	No change.
2:30:30	No change. The patient continues to stand on the standing board.

Figure 9

F. Gaps Between Chunks. You need not strain to fill a protocol with back-to-back chunks for the following reasons: (a) It might happen that a gap (time period between two chunks) will yield nonchunk tallies (Cs and Ss). (b) Sometimes, the record simply will not give enough information to infer anything between the end of one chunk and the beginning of another. (c) Sometimes, the 15-second rule for minimum length of chunks will produce gaps. Theoretically, however, chunking should be continuous and back-to-back, so your coding should ordinarily occur in this way. When a gap does occur in the coding, label it clearly between the brackets. (See Figure 10)

	5:37:30	The patient continues to eat.
	5:37:42	The patient puts down his fork and pushes the dinner tray away.
	5:37:51	The patient sits quietly now.
GAP		
	5:37:55	The patient reaches over to his bed and picks up the book there. He converses with NA James as he does this.
	5:38:12	The patient opens the book.

Figure 10

G. Chunks Involving Overt Idleness. There are many periods in the protocols during which the patients are behaviorally, overtly idle. Such occurrences should be marked as distinct chunks when and only when there is no other chunk occurring simultaneously. For example, Lying in Bed, or Sitting in Wheelchair, will be marked as a distinct chunk only when, judging from the protocol, that is all the target patient is doing. If he is also Watching TV or Reading, Lying in Bed is not to be marked off. In other words, chunks whose principal activity is some form of gross motor idleness or absence of other molar behavior should be kept as a pure case. They represent a kind of residual case, the times when the patient's behavior stream has damped down to a level near zero. Thus, if the patient is idle for an extended period, he then engages in some other nonidle behavior coded as a chunk, and then goes back to being idle, do not code the other chunk as overlapping one long idle chunk.

1. Sleeping vs. Other Idle Periods. Sleeping cannot be coded realiably, so you should never code a Sleeping chunk. Rather, code some other form of idle functioning; e.g., Lying in Bed, Lying on Mats.

2. Codes Accompanying Idle Chunks. Since the principal activity of idle chunks represents the pure case of the patient's doing nothing overtly, it would involve quite an intellectual strain to make inferences regarding O, I, and DI codes. Thus, for all idle chunks, the O code will be "300" (Irrelevant), the I code will be "300" (Irrelevant), and the DI code will be "4" (Irrelevant). Labels, settings, and beginning and ending time should be supplied.

H. Idle vs. Fidgeting, Diddling, etc. The system of chunk label classifications (which is applied later) includes a distinction among (a) Idle-Horizontal, (b) Idle-Upright, and (c) Fooling Around, Diddling, Gross Motor Fidgeting.

1. Idle-Horizontal and Idle-Upright will be used after the chunking process to classify the usual and frequent forms of idleness; e.g., Lying in Bed, Lying in Bed and Looking Out of Window, Waiting, Sitting in Wheelchair in Hall. Attach such descriptive labels to chunks in which the patient is essentially doing nothing in the gross motor sense, even if he shifts, changes position, or moves his limbs occasionally. Such movements are reasonable accompaniments of idleness. Attach your own descriptive labels and do not use "Idle-Horizontal" or "Idle-Upright."

2. If, and only if, the patient engages in relatively continuous and regular gross motor movements while located in one place, attach a chunk label that indicates something more than pure idleness; e.g., Engaging in Fidgeting Movements, Fidgeting, Fooling Around, Moving Back and Forth, Moving Up and Down.

I. Transfers, Transports, and Preparation for Them. If longer than 15 seconds, code all transfers (from chair to bed, mat to chair, etc.) and transports (room to PT, OT to PT, etc.) as distinct chunks, even when it appears arbitrary to do so, for two reasons: (a) We want transfers and transports in the major code system, with all the accompanying information. (b) This should take care of most instances of change of location. Transfers of less than 15 seconds should not be marked as chunks; when appropriate, give them S codes.

1. When marking transfers, include the preparation for transfer and cleaning up of transfer (finishing) in the chunk, primarily to avoid proliferation of tiny chunks, and partly because these are activities that are closely tied in terms of purpose and proximity in space and time. However, when marking transports, mark the activity involved in preparing for transport separately as a chunk if there is any such preparation, primarily because it will occur in one location and the transport will evolve and end elsewhere. Since the latter two involve an emerging separation in location, and since we will summarize by locations, mark them separately.

2. Be alert for transfers that obviously occurred while the observer's visual access to the patient was cut off, e.g., while a room door was closed. For example, if a patient was in his wheelchair when the door closed, and he was in bed when the door reopened, you can safely infer that a transfer took place. It is easy to overlook such events, so be on the lookout for them. When such a transfer takes place, and the period in question is four minutes or longer, give the transfer 2 minutes length; if the time in question is less than 4 minutes, assign the transfer one-half the total time. Whenever possible, use explicit information from the protocol to establish beginning and/or ending times. In the absence of such information, you should

usually use the beginning time of the period in question to anchor the beginning of the transfer. The only exception occurs when the period in question is immediately followed by a transport, but not preceded by a transport; in this case, anchor the ending time of the transfer at the end of the period in question. (See Figures 11 and 12)

Transport	1:04:25	The patient is being pushed in his wheelchair by NA James. They are moving down the hall.
	1:04:30	They enter the patient's room. NA Barnes follows them in.
Transfer	1:04:35	NA Barnes closes the door.
	1:05:00	No change. The door remains closed.
	1:06:00	No change.
Door Closed (Being Undressed and Put Into Bed)	1:06:35	
	1:07:00	No change.
	1:08:00	No change.
	1:09:00	No change.
	1:10:00	No change.
Lying in Bed	1:10:20	NA Barnes opens the door, and both NAs leave the room. The patient is lying in bed. The curtains are drawn and the room is dark. NA Barnes reports that the patient was undressed and put into bed.

Figure 11

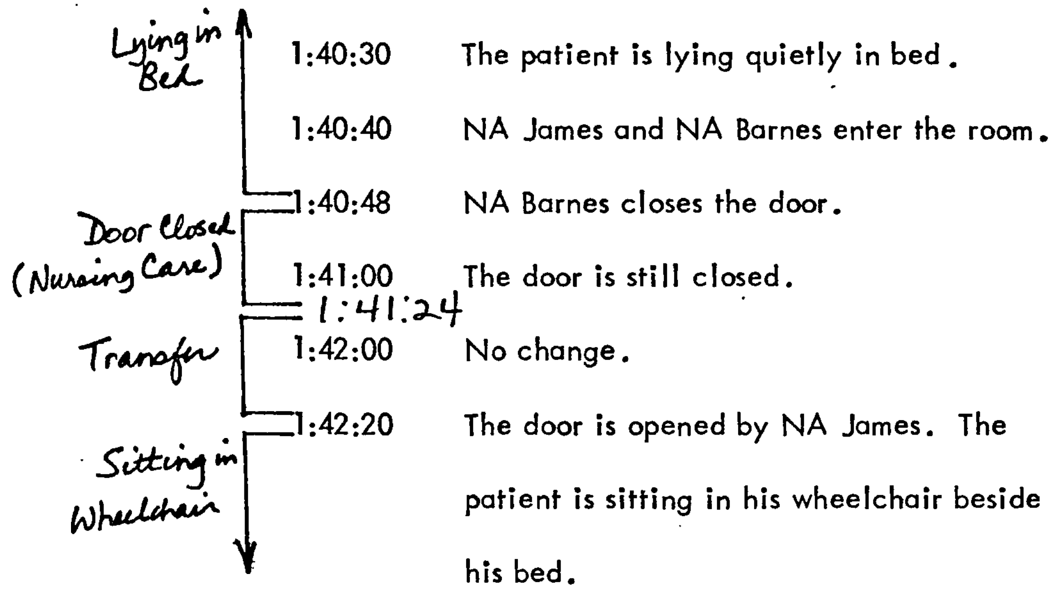


Figure 12

3. The system of chunk classifications to be applied later includes distinctions among Transporting Within a Setting, Transporting Between Settings, and Apparently Aimless Wandering Around and Exploring. Some movement events that might otherwise be chunked will probably be eliminated by the 15-second rule. One thing to remember when dealing with transports is that there might sometimes be more than one occurring in a movement sequence, e.g., Transporting to Concession Area, Sitting in Wheelchair, Transporting to OT. In any case, apply descriptive labels to movement chunks in such a manner that the chunks can be classified later. Do not use the chunk classification category labels.

J. Use of Standing Board and Progressive-Sitting Wheelchair. Since they are formal, prescribed treatments, Standing on Standing Board and Progressive Sitting may be chunks, with the possibility of all the accompanying codes (O, I, DI), even though the patient is ostensibly idle. Degree of involvement for Standing on Standing Board and Progressive Sitting will be 2 (passive).

1. Standing Board. A number of activities (e.g., checking blood pressure, other monitoring functions) are typical and necessary accompaniments of Standing on Standing Board and, therefore, should not be chunked separately or as overlapping. Furthermore, after the patient is lowered to the horizontal at the end of Standing Board, if someone checks his blood pressure or goes through a short monitoring sequence, include this in the Standing Board chunk if it is not separated from Standing Board by another chunk. If it is separated by another chunk, chunk it separately with a different chunk label.

2. Progressive Sitting. Two conditions are necessary for Progressive Sitting. (a) The patient must be sitting in an adjustable-back wheelchair, and (b) there must be no

clear indication in the protocol that the activity is not Progressive Sitting. In addition, at least two of the events in the following list must occur, together with the necessary conditions, in order to chunk Progressive Sitting.

(a) Monitoring of blood pressure by someone else

(b) Monitoring of pulse by someone else

(c) Adjusting the angle or tilt of back of wheelchair

(d) Inquiries by others regarding dizziness, nausea

(e) Tilting whole wheelchair backward for short periods

(f) Verbalizations about Progressive Sitting being in progress

If the activity is judged to be a Progressive Sitting chunk, do not code any of the events above as separate or overlapping, as they are typical and necessary accompaniments of Progressive Sitting. Progressive Sitting chunks begin when the patient has been transferred into the adjustable-back wheelchair even though conditions necessary for coding Progressive Sitting may not occur until much later in the protocol.

K. Using Telephone. In chunking the use of a telephone, deal only with the nonverbal aspects of the activity.

1. The cues for Using Telephone are generally prop-dependent. Using Telephone will usually begin when the patient comes in contact with telephone equipment and end when the equipment is removed.

2. A person who is directly involved in the physical act of handling and using the phone, e.g., someone who holds the receiver while the patient talks, will be coded as an O .

3. Do not code the person at the other end of the conversation as an O in the Telephoning chunk. If there is a mutual verbal interchange, indicate it with a C in the margin. If you can classify the other person (as when the patient says, "Hi, Mom"), use the appropriate three-digit code number with the C. In you cannot, use 300 (Cannot be Inferred).

4. If someone brings the phone to the patient and says, "Phone call for you," code that person for I.

L. Smoking. The activities involved in smoking cigarettes, cigars, etc., present special coding problems. Once the cigarette has been lit, the activity of smoking usually becomes quite intermittent. Let common sense guide you as you code this activity.

1. The act of getting out a cigarette and lighting it will usually be consistent enough to chunk and should be accompanied by appropriate O and I codes.

2. The individual puffs on a cigarette or cigar are usually too intermittent to chunk; do not strain to include them. If the patient holds the cigarette and takes the intermittent puffs himself, drop it. If someone else holds the cigarette while the patient takes a puff, give the other person an S.

M. Watching TV. To be coded as a chunk, TV Watching must be clearly described as a continuous activity, allowing for reasonable pauses and disruptions.

1. Persons who become involved in Watching TV by adjusting the TV, the speaker, or the prism glasses for the patient should be coded as O.

2. If no one else clearly instigates the chunk, apply 010 (patient himself) for I.

3. Because of the passive nature of Watching TV, you will usually code DI as being Level 2 (passive).

N. Listening to Radio. Just as the protocols do not include enough subtle information to discriminate between sleeping and other times when the patient is idle in bed, they do not include enough information to make reliable judgments of listening to the radio. Part of the problem here is that the auditory mode often involves no observable cues, such as looking or changing orientation. Even though the protocol will sometimes refer to listening to radio or snapping fingers in time to radio music, it is not possible to judge when listening begins and ends. Thus, do not chunk listening to radio. Rather, chunk whatever else the patient is doing.

O. Events Behind Closed Doors or Curtains. Normally, observers will have asked someone or otherwise given some specification of what occurs behind closed doors or curtains. Check through the protocol for information that may have been supplied after the event in question.

1. Whenever clearly defined and different activities occur in combination behind closed doors or curtains, without explicit reference to the beginning and ending times of each specific activity, make the sequence of activities one chunk and label the chunk with all of the activities that occur; e.g., Being Irrigated, Being Turned, Being Dressed, and Taking Pills.

2. When no specification is given whatsoever, use the general context to make a reasonable assumption for chunking. In the ward, with ward staff behind the door or curtains, you will usually chunk Receiving General Patient Care, or Receiving Undetermined Patient Care. Under these circumstances, be sure to watch for transfers that obviously occur.

3. The beginning time for chunks behind closed curtains or doors will usually be the point at which the curtains or doors are closed around both the staff and patient. The ending time will usually be the point at which the curtains or doors are opened again. Be sure to examine the text closely; use the most reasonable cues possible to determine the beginning and ending times of these chunks.

P. Occasional Dual Coding of Verbalizations. Sometimes, usually in formal, supervised exercises, verbal input and exchange is an essential accompaniment of the principal activity; i.e., the activity is organized and sequenced properly by essential verbal inputs and verbal feedback.

1. In those rare cases in which such verbal staging, organizing, sequencing, and pacing of a principal activity is the only described involvement by someone other than the target patient, you may use the verbal involvement to code that person as an O. Again, the requirement is that a person be an essential component of the principal activity of the chunk.

2. Ordinarily, the use of such verbalizations to code O should be restricted to formal exercises and then only if the protocol says clearly that the verbalizations were used to organize, pace, and sequence the principal activity. Thus, if the protocol says only, "They exchange a few comments," do not use this description of verbalizations to code an O.

3. If the protocol indicates that verbalizations leading you to code an O also involved mutual interchange, code a marginal C as well as O.

Q. When the Patient is Inaccessible for Observation. If a patient is in the hospital, but behind a closed door, he is, by convention, accessible for observation, and the protocol should be coded in routine fashion. Occasionally, however, a patient will be gone from the hospital during all or part of an observation period. As long as the patient is accessible to the observer (e.g., in a car in the hospital driveway), his activities may be chunked. When the patient is not accessible, his activities may not be chunked, even though coding information may be available retrospectively. The entire period in which the patient is inaccessible is coded only with label, begin time, and end time.

APPENDIX B

Shortform Coding Sheet

TIME	ACTIVITY	OTHER	I	DI	C	locat

APPENDIX C

Guidelines for Observers
and
Guidelines for Coders

Guidelines for Observers

1. As in the narrative system, your focus is on what the target patient is doing.
2. Record enough of the context of the patient's behavior to make it intelligible.
3. Every half-minute, ask yourself, "What did the patient do most during the preceding half-minute?" Record the activity that lasted longest. Leave out anything that lasts less than 15 seconds.
4. Every half minute, ask yourself, "Was anyone else directly involved in the activity I just recorded?" Use the guidelines offered for coders in the narrative system. If someone else was directly involved (for <u>any</u> length of time) in the half-minute, record that person's name. If no one was involved in that half-minute, write "none." Make this decision every half-minute, <u>independently</u> of any decisions that preceded it.
5. If the activity you have recorded is different from the one that preceded it, ask yourself, "Who instigated this?" Use the guidelines offered for coders in the narrative system. Record the person's name. If the activity you recorded is a continuation of the preceding entries, do not denote an instigator. Make this decision for every new activity you record.
6. Every half-minute, ask yourself, "What has been the patient's general degree of involvement in the activity I have recorded?" Use the guidelines offered for coders in the narrative system. Make this decision every half-minute, <u>independently</u> of any decision that preceded it.
7. Every half-minute, ask yourself, "Has anyone conversed with the patient?" Use the guidelines offered for coders in the narrative system. Record the name of everyone who conversed with the patient in that half-minute, even if the conversation did not occur during the activity you have recorded. If no one conversed with the target patient during the half-minute, write "none." Make this decision

every half-minute, independently of any decisions that preceded it.

8. In the first half-minute of every page of your observation, record the location of the target patient. You need not record location again unless it changes. If the location changes, write the new location in the half-minute in which it is entered.

9. Please write neatly. If you are uncertain whether to include or omit some information, do include it--but never write so much information in any cell that you write below the bottom line.

Guidelines for Coders

1. As in the narrative system, your focus in on what the target patient is doing.
2. Mark off the observation into chunks. Use the guidelines offered in the narrative system.
3. Give each chunk a descriptive label.
4. Write in the appropriate code numbers for all the different persons listed as "other" in the chunks, using the pre-set codes from the narrative system.
5. Write in the code number for instigator(s), using the codes from the narrative system.
6. Write in the modal code for the target's degree of involvement.
7. Write in the code numbers for persons who conversed with the target patient for each chunk, using the codes from the narrative system.
8. Write in the code number for every location listed using the codes from the narrative system.
9. Never make assumptions. If a cell is empty, ask the observer about it. If you do ask questions, write the answers down in the protocol so the checker won't have to ask questions too. Always ask the observer to check your coding as you proceed through the protocol.
10. Transfer the coded data to the chunk data format, completing one chunk at a time. If a chunk has no conversation codes, mark a null sign ($\emptyset$) for the C-code.
11. Write neatly.
12. Check your work.