

1989

# A study to test the reliability of an independent living skills assessment instrument

Michelle Martin Mueller  
*San Jose State University*

Follow this and additional works at: [https://scholarworks.sjsu.edu/etd\\_theses](https://scholarworks.sjsu.edu/etd_theses)

---

## Recommended Citation

Mueller, Michelle Martin, "A study to test the reliability of an independent living skills assessment instrument" (1989). *Master's Theses*. 3113.

DOI: <https://doi.org/10.31979/etd.rk92-htn4>

[https://scholarworks.sjsu.edu/etd\\_theses/3113](https://scholarworks.sjsu.edu/etd_theses/3113)

This Thesis is brought to you for free and open access by the Master's Theses and Graduate Research at SJSU ScholarWorks. It has been accepted for inclusion in Master's Theses by an authorized administrator of SJSU ScholarWorks. For more information, please contact [scholarworks@sjsu.edu](mailto:scholarworks@sjsu.edu).

## INFORMATION TO USERS

The most advanced technology has been used to photograph and reproduce this manuscript from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book. These are also available as one exposure on a standard 35mm slide or as a 17" x 23" black and white photographic print for an additional charge.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

# U·M·I

University Microfilms International  
A Bell & Howell Information Company  
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA  
313/761-4700 800/521-0600



Order Number 1337847

**A study to test the reliability of an Independent Living Skills  
Assessment Instrument**

Mueller, Michelle Martin, M.S.  
Stack-Dunne, Susan Elizabeth, M.S.

San Jose State University, 1989

**U·M·I**  
300 N. Zeeb Rd.  
Ann Arbor, MI 48106



A STUDY TO TEST THE RELIABILITY OF  
AN INDEPENDENT LIVING SKILLS ASSESSMENT INSTRUMENT

A Thesis

Presented to

The Faculty of the Department of Occupational Therapy  
San Jose State University

In Partial Fulfillment

of the Requirements for the Degree  
Master of Science

By

Michelle Martin Mueller

and

Susan Elizabeth Stack-Dunne

May 1989

APPROVED FOR THE DEPARTMENT OF OCCUPATIONAL THERAPY

*Morag Paterson*

Morag A. Paterson, M.A., OTR

*Lela A. Llorens*

Lela A Llorens, Ph.D., OTR, FAOTA

*Denise O. Hewitt, OTR*

Denise O. Hewitt, OTR

APPROVED FOR THE UNIVERSITY

*Serena N. Stanford*

## Abstract

A study of test-retest and interrater reliability was conducted on the Independent Living Skills Assessment Instrument, a functional capacity evaluation tool which is used at the Center for Independence of the Disabled in Belmont, California. The instrument measures a comprehensive range of functional living skills and was administered by registered occupational therapists. Persons with permanent physical disabilities who were living in the community were evaluated. The test-retest reliability portion of the study consisted of two administrations of the instrument by the same rater to each of five subjects. The interrater reliability test involved three subjects, all registered occupational therapists, who administered the instrument one time to each of ten participants. Overall test-retest reliability of the instrument was 90.3% at the 90% level of confidence. Overall interrater reliability was 61.3% at the 90% confidence level.



## Acknowledgements

The authors wish to thank the members of their committee for their ongoing review and recommendations: Morag Paterson, Dr. Lela Llorens, and Denise Oosterman Hewitt. A special thanks goes to Lela Llorens, for her meticulous review of drafts and invaluable input to content and format.

We would like to thank our statisticians: Alex Leung, who assisted with the study design; and Anirudh Kulkarni, who provided assistance with the statistical analysis. Thanks also goes to Paul Tierney, who assisted with the printing.

We would also like to thank the staff and clients of the Center for Independence of the Disabled; in particular we thank Cathy Heumann for sharing her clinical expertise, and Amy O'Connor and Dori Wheeler for their administrative support. Thanks also goes to Lois Glover for her assistance with data collection.

Lastly, we thank our families: Dick and Casey Mueller, Mary Stack-Dunne, Rachel, Stuart, Rebecca, and Katie Kopperman for their support and patience through the process of completing this thesis.

## Table of Contents

Abstract . . . . .	iii
Acknowledgements . . . . .	iv
List of Tables . . . . .	vii
CHAPTER 1. INTRODUCTION . . . . .	1
Purpose . . . . .	1
Background . . . . .	1
Statement of the Problem . . . . .	3
Objectives, Questions, and Hypotheses . . . . .	5
Definitions . . . . .	7
Assumptions . . . . .	8
Limitations . . . . .	8
Significance of the Study . . . . .	9
CHAPTER 2. REVIEW OF THE LITERATURE . . . . .	10
Occupational Therapy and Independent Living . . . . .	11
Instrumentation to Measure Independent Living Skills . . . . .	14
Summary . . . . .	25
CHAPTER 3. RESEARCH PROCEDURES . . . . .	28
Objectives . . . . .	28
Questions . . . . .	29
Null Hypotheses . . . . .	30
Participants for the Study . . . . .	30
Data Collection Techniques for Test-Retest Reliability . . . . .	34
Data Collection Techniques for Interrater Reliability . . . . .	36

CHAPTER 4. ANALYSIS OF DATA . . . . .	40
Purpose . . . . .	40
Analysis Approach . . . . .	42
Results for Test-Retest Reliability . . . . .	44
Results for Interrater Reliability . . . . .	50
CHAPTER 5. RESULTS, IMPLICATIONS, SUMMARY, AND	
RECOMMENDATIONS . . . . .	60
Discussion . . . . .	60
Implications for the Profession . . . . .	68
Summary . . . . .	70
Recommendations . . . . .	72
References . . . . .	74
Appendix A . . . . .	77
Appendix B . . . . .	98
Appendix C . . . . .	100

## List of Tables

1. Demographic Data for Test-Retest Client Sample . . . . .	32
2. Demographic Data for Interrater Reliability Client Sample . . . . .	33
3. Order of Administration of Independent Living Skills Assessment Instrument for Interrater Reliability . . . . .	37
4. Number of Discrepancies for Test-Retest Reliability . . . . .	45
5. Predicted Test-Retest Reliability by Section . . . . .	47
6. Percent Discrepancy for Test-Retest Reliability . . . . .	48
7. Number of Discrepancies for Interrater Reliability . . . . .	51
8. Predicted Interrater Reliability by Section . . . . .	53
9. Percent Discrepancy for Interrater Reliability for Each Client . . . . .	54
10. Mean Percent Discrepancy for Interrater Reliability . . . . .	55
11. Pairwise Comparisons of Sections Using the Tukey Method . . . . .	58

## CHAPTER 1

### INTRODUCTION

#### Purpose

The purpose of the study was to determine the reliability of the Independent Living Skills Assessment Instrument (ILSAI).

#### Background

The Center for Independence of the Disabled (CID) in Belmont, California, was opened in October, 1979, by a group of disabled people. Their purpose was to "provide peer-oriented support and services to disabled adults living in San Mateo County and to advocate for the rights of disabled people everywhere" (CID, p. 1). This philosophy reflected that of the independent living (IL) movement and also served as the foundation for other independent living centers throughout the country.

The services that CID provided included attendant care referral, housing referral, peer counseling, accessibility modifications, information and referral, and two independent living skills programs.

One of the latter programs served clients on a one-to-one basis, usually in the home or work environment. The other program was designed to target young adults between the ages of 16 and 30 who have physical, sensory, or developmental disabilities. Both programs were coordinated by registered occupational therapists.

When the first independent living skills program was opened, the occupational therapist recognized the need to have an assessment instrument which would accurately measure the functional abilities of the clients and was also modeled on the IL philosophy. Such an assessment instrument would not be designed to compare individuals to each other, but rather to gather baseline data on an individual's functional ability for comparison with subsequent performance following intervention. This instrument is called the Independent Living Skills Assessment Instrument (ILSAI) (see Appendix A).

There are two features of the ILSAI which reflect the IL philosophy. The first is the inclusion of categories that allow for the choice of the client to utilize assistance to perform an activity when performing it independently would be impractical. The second is the method used to rate the level of independence of the client. If the client utilizes personal assistance, by choice or out of necessity, and is able to direct the assistant in carrying out this activity, this is scored as "independent" in that activity. The rationale for using this method

of ranking is to acknowledge that independence can be defined both as physical ability, as well as the cognitive ability to direct an attendant. (A list of the ranking of each graded category can be found in Appendix A.)

The ILSAI has been revised on two subsequent occasions. The latter revision included the creation of a formal protocol that specified the criteria by which each task was evaluated (see Appendix A). When the protocol for administration was introduced into the assessment process it then became possible to conduct reliability studies on the ILSAI.

### Statement of the Problem

"Occupational Therapy is the study of human occupations...and the management of the adaptive behavior required to perform these occupational functions" (Reed and Sanderson, 1980, p. 5). The practice of occupational therapy, as presented by Reed and Sanderson's Process Model, includes assessment, treatment planning, treatment, reassessment, and analysis. The efficacy of treatment, therefore, rests on a foundation of comprehensive and accurate assessment. Accurate assessment is dependent upon reliable instruments.

There are insufficient assessment instruments available in occupational therapy that have been developed or tested in a formal, scientific manner. Often constraints of time, money, and expertise have led professionals to develop and use instruments that have not been sufficiently tested. Occupational therapists working in a variety of settings have been hampered by these constraints. A preliminary study on an instrument to assess independent living skills was needed both in the field of occupational therapy and by CID.

Recognizing the need to have research conducted, CID approached the investigators and requested that a reliability study be carried out on the ILSAI, which was being administered to approximately 60 clients annually. In addition to the clinical applications at CID, it was believed that this assessment instrument also had the potential to be of value to occupational therapists in other settings who were evaluating similar functional skills. In order to be of substantive value to other therapists, however, the assessment instrument needed to be subjected to tests of reliability.

Internal consistency, or test-retest measures, are fundamental to the reliability of an assessment instrument. In addition, CID was concerned about the level of consistency among raters. It was for these reasons that these two particular reliability tests were included in the study.



## Objectives, Questions, and Hypotheses

This study sought to determine test-retest and interrater reliability.

### Objectives

The objectives of the study were to determine the:

1. Test-retest reliability of the Independent Living Skills Assessment Instrument.
2. Interrater reliability of the Independent Living Skills Assessment Instrument.

### Questions

The questions to be answered by the research were:

1. Is there a significant level of test-retest reliability for each section of the Independent Living Skills Assessment Instrument?
2. Which section(s) of the Independent Living Skills Assessment Instrument has/have the highest level of test-retest reliability?
3. Which section(s) of the Independent Living Skills Assessment Instrument has/have the lowest level of test-retest reliability?

4. Does the Independent Living Skills Assessment Instrument as a whole have a significant level of test-retest reliability?

5. Is there a significant level of agreement among three raters for each section on the Independent Living Skills Assessment Instrument?

6. Which section(s) of the Independent Living Skills Assessment Instrument has/have the highest level of agreement among three raters?

7. Which section(s) of the Independent Living Skills Assessment Instrument has/have the lowest level of agreement among three raters?

8. Does the Independent Living Skills Assessment Instrument as a whole have a significant level of agreement among three raters?

#### Null Hypotheses

1. There will be no significant level of test-retest reliability on the Independent Living Skills Assessment Instrument.

2. There will be no significant level of interrater reliability among three raters of the Independent Living Skills Assessment Instrument.

## Definitions

Terms used within this study were defined as follows:

**DISABLED:** having some physical and/or sensory impairment which significantly affects functional performance of independent living skills.

**INDEPENDENT:** free from the influence, guidance, or control of another or others; self-reliant.

**INDEPENDENT LIVING CENTER:** consumer- and community-based organization which provides support services to disabled individuals in the community in order to maintain or increase independence.

**INDEPENDENT LIVING SKILLS:** those functional abilities that are necessary for self-sufficiency in a living situation; in this study, these are considered to be synonymous with activities of daily living (ADL).

**INDEPENDENT LIVING SKILLS ASSESSMENT INSTRUMENT:** a tool designed to evaluate the functional skills level of disabled adults who are living in the community.

**INTER-RATER RELIABILITY:** the quality of consistency of an instrument among raters.

**TEST-RETEST RELIABILITY:** the quality of consistency of an instrument when repeated over time with the same or similar population.

## Assumptions

The following were the assumptions held for this study:

1. Clients did not learn independent living skills from their participation in the assessment process.
2. Clients provided accurate information when responding to items during administration of the Independent Living Skills Assessment Instrument.
3. Differences in the participants' responses were not due to a change in their level of functioning.
4. Given the requirements stated in the methodology, the raters were competent to administer the Independent Living Skills Assessment Instrument.
5. The Independent Living Skills Assessment Instrument has a high degree of face validity.

## Limitations

The following limitations were identified for this study:

1. The scores may have been affected by clients' knowledge that they were participating in a research project.

2. The study sample was small and was not randomly selected; the research results, therefore, cannot be generalized to a wider population.

### Significance of the Study

As medical technology becomes increasingly advanced, the population of disabled persons increases. In the expanding field of rehabilitation, habilitation, and post-hospitalization services, the occupational therapist is qualified to evaluate and teach functional living skills. The results of this study may provide evidence of the reliability of an assessment instrument which could be used by occupational therapists in a variety of settings or it may provide evidence of areas that need further testing. On a broader level, the study will contribute to the growing body of knowledge concerning instrument development within the field of occupational therapy.

## CHAPTER 2

### REVIEW OF THE LITERATURE

The purpose of the literature review was twofold. First, it was to assay the relationships among occupational therapy, the Independent Living (IL) movement, and activities of daily living (ADL). This section includes a review of occupational therapy's standards of practice related to ADL, as well as background on the IL movement in the United States and its relationship to the theory and practice of occupational therapy. The second purpose was to survey current literature regarding assessment instrumentation, particularly that regarding assessment instruments which are similar to the Independent Living Skills Assessment Instrument (ILSAI) and have been studied for the purpose of evaluating reliability. Special attention has been given to the review of those assessment instruments which have been tested for test-retest and/or interrater reliability.

## Occupational Therapy and Independent Living

Documents published by the American Occupational Therapy Association (AOTA) specify that activities of daily living are a primary dimension of practice. According to the document, "Occupational Therapy: Its Definitions and Functions" (AOTA, 1972), occupational therapy programs are divided into three categories: prevention and health maintenance programs, remedial programs, and daily life tasks and vocational adjustment programs. "Entry Level Role Delineation for OTRs [registered occupational therapists] and COTAs [certified occupational therapy assistants]" (AOTA, 1981) lists four areas of practice for which the occupational therapist is responsible. The first of these four categories is "Independent Living/Daily Living Skills" (p. 5). The authors of both documents define activities of daily living as an integral part of the practice of occupational therapy.

In a short paper on occupational therapy's role in independent living, Baum (1980) stated that "independent living is a concept occupational therapists have long valued" (p. 773). She pointed out that occupational therapy as a profession has traditionally supported the concept of independent living as a primary goal of treatment. The concept of independent living, although discussed and included as part of occupational therapy's theory and practice, has moved

from "concept" to "movement" as it has expanded from a professionally-oriented concern to a consumer-based viewpoint.

DeJong (1979) identified the core constituency of the IL movement as individuals who are physically disabled. He noted that there are additional groups that are an important part of the movement, such as professionals who work with people with disabilities and special interest organizations. Levels of involvement of these groups within the IL movement vary greatly. DeJong also discussed the concept of a paradigm and its relationship to the IL movement. He defined a paradigm as a framework of thinking "by which problems are identified and solved. A paradigm also prescribes the technology needed to solve a given problem" (p. 442). He stated that a paradigm shift occurs when events cannot be explained by a current framework of thinking.

According to DeJong (1979), the dominant paradigm in disability today is that of rehabilitation. He stated that:

in the rehabilitation paradigm, problems are generally defined in terms of inadequate performance in ADL or in terms of inadequate preparation for gainful employment. In both instances, the problem is assumed to reside in the individual. It is the individual who needs to be changed. To overcome his/her problem, the disabled individual is expected to yield to the advice and instruction of a physician, physical therapist, occupational therapist, or a vocational rehabilitation counselor. The disabled individual is expected to assume the role of "patient" or "client." While the goal of the rehabilitation process is maximum physical



functioning or gainful employment, success in rehabilitation is to a large degree determined by whether the patient or client complied with the prescribed therapeutic regime. (p. 442)

In contrast, DeJong wrote that:

according to the IL paradigm, the problem does not reside in the individual but often in the solution offered by the rehabilitation paradigm--the dependency-inducing features of the physician-patient or professional-client relationship. Rehabilitation is seen as part of the problem, not the solution. The locus of the problem is not the individual but the environment that includes not only the rehabilitation process but also the physical environment and the social control mechanisms in society-at-large. To cope with these environmental barriers, the disabled person must shed the patient or client role for the consumer role. Advocacy, peer counseling, self-help, consumer control, and barrier removal are the trademarks of the IL paradigm. (p. 443)

Verville (1979) detailed relevant legislative history and its effects on the IL movement over the past 60 years. He emphasized the effect of 1978 legislation, known as Title VII or Comprehensive Services for Independent Living. This legislation allowed services for the disabled to be expanded to include those individuals who were not vocationally-oriented and might never be employed. Because Verville selected legislative items that spanned a relatively long period of time and were specific to disability, his article gave a clear picture of the legislative changes pertinent to disability. Verville's article supports and complements that of DeJong.

In 1981, the AOTA Commission on Practice authored an official position paper which outlined the role of occupational therapy in independent living. This paper was significant in that the position taken paralleled the values presented in DeJong's IL paradigm. Most importantly, the AOTA supported the client as the center of control and decision-making in the client-therapist relationship, a concept which is central to the IL movement philosophy. This perspective is in contrast to that of the rehabilitation paradigm, which supported the therapist as decision- and goal-maker for the patient or client. Although occupational therapy is traditionally viewed from the rehabilitation paradigm, the AOTA has taken the position of valuing the client-centered approach to the provision of services.

In conclusion, the philosophies and practices of occupational therapy and the IL movement in the area of activities of daily living overlap and support each other. The IL movement has a special significance to those occupational therapists who work in settings where evaluation and treatment directly affect clients living in the community.

#### Instrumentation to Measure Independent Living Skills

Current literature related to the standardization of assessment instruments was reviewed. Anastasi (1976) defined standardization

as "uniformity of procedure in administering and scoring the test" (p. 25). According to Keith (1984), "results that can be generalized from one patient to another or from one facility to another requires [sic] standardized measures" (p. 76). Recognizing the importance of developing standardization procedures for assessment instruments which are used in clinical settings, a number of investigators have carried out standardization studies.

Only those assessment instruments which were similar to the ILSAI were included in this review. The instruments which the authors considered similar for the purposes of this study had major components in ADL, or were functional life or IL scales. These were considered to be similar because the content and focus of the instruments were comparable to those of the ILSAI. There was no evidence found in the literature that an assessment instrument had been developed and standardized, in whole or part, that targeted both the same population as the ILSAI and was based on the IL philosophy.

Jette (1980) performed studies designed to evaluate both repeatability and interrater reliability on an instrument similar to the ILSAI. The Functional Status Index was designed to measure degree of dependence, degree of difficulty, and the amount of pain experienced in performing specific activities of daily living. The study used a sample of 149 adults with rheumatoid arthritis. Scores

of test-retest and interobserver reliabilities were calculated by category rather than item by item. Categories were analyzed statistically using an intraclass correlation coefficient. In general, individual measures of functional dependence achieved test-retest and interobserver reliabilities of .6 or greater. The measures of degree of difficulty in performing activities in all five functional categories were found at test-retest and interobserver reliability levels of .7 and greater, with a few exceptions. Both measures for functional pain were .6 or greater. Exceptions were noted and tables given for further illustration. The author noted that:

a critical problem in assessing degree of test-retest and interobserver and intertime reliability is selecting an interval long enough to minimize the potential confounding influence of memory but short enough to minimize the possibility of real changes in the function being assessed. (p. 398)

Discussion of the results of the study, including limitations, was clear and credible. One potential additional limitation, though, was not addressed; 65 respondents were included in the test-retest process. However, 26 respondents were interviewed the second time by the original interviewer and 39 respondents were interviewed the second time by a different person. No rationale was given for why this procedure was used. The use of two different interviewers

may represent a confounding variable and should have been included as a limitation of the study.

Charlton, Patrick, and Peach (1983) conducted a study in England which included reliability measures as part of a larger study on disability. The instrument used for the test-retest portion of the study was called the functional limitations profile (FLP). (The FLP was a modified version of the sickness impact profile, an instrument that was developed in the United States and contained 136 yes/no statements concerning restrictions on activity.) The repeatability study used a sample of 30 disabled patients attending a health clinic. Doctors were asked to choose patients with conditions that were unlikely to be resolved within 48 hours. Thirty of these patients were asked to complete a questionnaire twice, first at the clinic and then 48 hours later at home. They were instructed to return the second one to the investigators by mail. Item agreement was measured by noting the frequency of the endorsement of each item on both questionnaires and then stating that number as a percentage.

The authors did not describe how the subjects were chosen from the doctors' recommended list. Also, they did not state the percentage of return of the mail-in questionnaire. According to the table that listed the results, the investigators received a 100 percent return, an unusually high rate of return for a self-administered mail-in questionnaire. If the authors included only those 30 respondents

who returned both of the questionnaires, then the results of the study would be based on a biased sample. In addition, the discussion of the results was general and somewhat inconclusive. The authors claimed that the sample was too small to estimate the reliability of any one item accurately and the reliability for repeatability was reported as "low" with the range of item agreement varying from 22% to 100%.

McLaughlin (1980) carried out a reliability study in conjunction with developing an instrument that measured self-care behaviors of persons with multiple sclerosis. She designed the study in response to a need to establish a systematic measurement of chronic illness. In order to measure reliability, McLaughlin developed a closed-ended questionnaire, which was mailed to 20 persons. These persons constituted a convenient sample selected by age, sex, and degree of disability. The subjects were instructed to fill out the questionnaire as well as give written feedback on clarity and format. Two weeks later this process was repeated with the same subjects. Following a revision of the questionnaire the study was repeated with a new sample of 20 persons. The Pearson product-moment correlation measure was used to obtain a coefficient of .79. Because McLaughlin could not find a statistical value for a behavior inventory similar to the one she had developed, she compared her outcome with a median value for attitude scales, which was .79. McLaughlin's

reliability measure is then "considered a good indication of its reliability" (p. 113).

McLaughlin's (1980) study was carefully designed and the author stated clearly her reasons for choosing to study certain measures, the meaning of specific outcomes, and the way each one affected the development of the instrument. Her statistical procedures also appeared to be carefully selected and analyzed. Inclusion of tables within the text to support the statistical analysis would have made the results more clear.

Klein and Bell (1982) developed and tested the Klein-Bell ADL scale, which is designed to allow the disabled person to use a variety of methods to achieve the desired behavior. The scale includes "critical and easily observable components of ADL behavior of virtually all people, handicapped or able-bodied" (p. 336), and is rated by the observer. The ratings are considered equal whether or not assistive devices are used. The authors obtained a score for interrater reliability by rating 20 subjects independently, using pairs of occupational therapists or rehabilitation nurses. A total of three pairs of therapists and three pairs of nurses participated in the study. The authors reported that the Klein-Bell scale received 92% agreement among raters.

Although the authors noted that the evaluators were not extensively trained, they did not specify the method of training used.

No other details were given to describe how the reliability measure was obtained. Although the result suggested a high percentage of agreement for the Klein-Bell scale, the authors stated that this was a conservative figure. They assumed that with extensive training the percentage of agreement would be even higher. It was difficult to give credence to this assertion without more detailed information on the methodology, including the extent and type of training that was given to the evaluators.

Sparrow and Cicchetti (1978) assessed interrater reliability by using the observed behavior of a sample of institutionalized mentally retarded children. The scale was designed as a behavior rating inventory that included self-help skills. Total observer-reliability scale scores between two raters were .77 using a standard Pearson product-moment correlation, with a Kendall's Tau statistic of .70. The average percentage agreement was 88%. The data for this reliability study were based upon the 63 out of 68 items that were independently scored by the two observers.

The procedures used by Sparrow and Cicchetti (1978) were not clearly written, so it was difficult to determine the methodology of the study. It would seem that the professional staff who observed the behavior were then interviewed by psychologists, who recorded the information on the scale. Using this method to record the data may have affected the results, but the authors did not justify their



choice of methodology, nor did they mention its potential effects on the results.

Kuriansky and Gurland (1976) described "the background, rationale, usefulness, and specific administration and scoring procedures" (p. 343) of a performance test designed to measure the self-care capacity of geriatric psychiatric patients. The authors claimed that "reliability based on independent ratings of the interviewer and the observer was good (i.e., .902)" (p.348). It is not stated whether this number represents a percentage of agreement or a correlational measure. The research procedures were not stated clearly in the article and the results were not supported with a presentation or summary of any raw data.

Wolber and Lira (1981), in their study of a "Relationship between Bender Designs and Basic Living Skills of Geriatric Patients," generalized an interrater reliability coefficient of .92 for the Basic Living Skills assessment scale. The research report was relatively short and well-organized, but did not describe how this measure was reached or the methods that were used.

Brorsson and Asberg (1984) reported a study of interrater reliability and validity of the Katz Index of Independence in ADL completed at a department of internal medicine in Sweden. Developed and standardized in the United States more than 20 years earlier, this instrument for measuring the level of functioning in

chronically ill and aging populations was tested in Scandinavia so it could be recommended for use there. The instrument consists of six activities of daily living which are observed and recorded on an eight-point scale of independence. One hundred subjects were chosen by a resident physician, 50 from each of two wards in a county hospital. Subjects were at least 55 years old and had been on the ward for a minimum of three days. Most patients had more than one medical diagnosis. Nurses trained in the use of the instrument observed the patients as part of their regular daily work and rated them independently. The nurses were "instructed to note down the most dependent performance during the last 24 hours" (p. 127).

Results of interrater reliability reported for the study were scant. Brorsson and Asberg (1984) reported only that "the inter-observer variability was low" (p. 127). While considerable detail was given with regard to other studies that were carried out on the instrument at the same time, the authors failed to give any data to support this statement. In addition, methods used to select and observe subjects were questionable. The authors reported that the residents chose patients as subjects "stratified according to a subjective estimation of need for care to avoid merely totally independent or totally dependent patients" (p. 126). No reasons were given for the intended exclusion of patients on either end of the spectrum or the methods used to evaluate the patients for this potentially biased

sample. The authors also stated that "the observation task was carried out as part of the regular work" (p. 126) of the nurses. Since, as the authors pointed out in the discussion, "in nursing care there is a tendency to help patients more than they actually need" (p. 130), this method of data collection during the nurses' regular work hours may have interfered with the results.

Watts, Kielhofner, Bauer, Gregory and Valentine (1986) have reported the development of the Assessment of Occupational Functioning (AOF). The AOF was designed as a screening tool to "assess the functional capacity of residents in long term treatment settings who have physical and/or psychiatric problems" (p. 231). Development of the AOF included studies in both test-retest reliability and interrater reliability. The development of the instrument was based on the Model of Human Occupation which "conceptualizes humans as open systems interacting with their environments" (p. 235). The evaluation includes three large categories (or subsystems): the volition, the habituation, and the performance systems. There are six total categories in the evaluation, each of these being a part of one of the larger subsystems: values, personal causation, interests, roles, habits, and skills. The evaluation has two components: a semistructured interview and a rating of each of the six categories on a five-point scale from "absent" to "fully adaptive" (p. 235).

A small pilot test for interrater reliability was completed on 13 subjects using two occupational therapists as raters. A Pearson product-moment correlation of .85 was achieved. Further revisions were made and a larger study was carried out with 83 subjects from two medical institutions and the community. All subjects were 60 years old or older and institutionalized. All had a psychiatric diagnosis and some also had additional physical diagnoses. The interrater reliability portion of the study was conducted only with inpatients. Therapists conducted an audiotaped interview with subjects in the inpatient facilities. These audiotapes were then reviewed by the interviewer and two other researchers and the AOF instrument was completed for each subject. For purposes of the repeatability study, 14-to-21 days later, the audiotaped interview was readministered to the institutionalized subjects and evaluated in the same manner.

Both test-retest reliability and interrater reliability were considered by the authors to be "above minimal acceptable levels" (p. 238), although scores on some individual items were below acceptable levels. Pearson product-moment correlations calculated for test-retest reliability, calculated separately for different locations in which the study was carried out, showed that total score Pearson product-moment correlations ranged from .70 to .90, which was greater than the authors' cited minimal acceptable correlation of .60.

Interrater reliability was determined using an intraclass correlation coefficient for individual items and for the total score of all institutionalized subjects, which at .78, was considered by the authors to be a "substantial" correlation. No minimum acceptable score was reported for this portion of the analysis. Coefficients for individual items in both tests of reliability varied and in some cases fell below the acceptable levels. Discussion of the shortcomings of several individual items and their use in the instrument was included in the article.

This article demonstrated an excellent discussion of a well-organized attempt to develop and refine an occupational therapy assessment instrument for a specific screening purpose. The process and implementation, as well as the description and discussion of the AOF, were organized and comprehensible. Limitations of the study and its results were included throughout the paper and discussion of the instrument's potential use as a reliable and useful instrument was credible.

### Summary

No evidence was found in the current literature to indicate that an assessment instrument which targets physically disabled persons who are living in the community has been developed and wholly or

partially tested for reliability. Attempts have been made, however, to standardize assessment tools which these authors considered similar enough for the purposes of this review. The researchers used a variety of methodologies. Jette (1980) and Watts, Kielhofner, Bauer, Gregory, and Valentine (1986) conducted studies of test-retest and interrater reliability on instruments designed specifically for use by occupational therapists. Charlton, Patrick, and Peach (1983), and McLaughlin (1981) measured repeatability on an instrument by mailing self-administered questionnaires to the subjects.

The interrater reliability studies were carried out using a variety of methodologies. Klein and Bell (1982) used pairs of raters to evaluate the subjects. Sparrow and Cicchetti (1978) used a similar methodology, but extended the process by having a psychologist record the ratings of the observers. Kuriansky and Gurland (1982) had one evaluator interview the subject, while a second evaluator observed the interview. Brorsson and Asberg (1984) had pairs of nurses independently evaluate the subjects during their regular daily work. Wolber and Lira (1981) did not provide enough information to give a clear understanding of the research procedures used in their study. Many of the authors failed to give sufficient detail on the methodology of their studies. Most, however, offered acceptable statistics for test-retest and interrater reliability and stand as an initial, credible attempt to standardize instruments for

use in the health care field.

## CHAPTER 3

### RESEARCH PROCEDURES

The authors sought to determine test-retest and interrater reliability of the Independent Living Skills Assessment Instrument (ILSAI). The study was set up in two parts: the first part examined test-retest reliability and one rater administered the instrument twice to each of five participants; the second test involved interrater reliability and in this case three raters administered the instrument to each of ten participants.

#### Objectives

The objectives of the study were to determine the:

1. Test-retest reliability of the Independent Living Skills Assessment Instrument.
2. Interrater reliability of the Independent Living Skills Assessment Instrument.



## Questions

The questions to be answered by the research were:

1. Is there a significant level of test-retest reliability for each section of the Independent Living Skills Assessment Instrument?
2. Which section(s) of the Independent Living Skills Assessment Instrument has/have the highest level of test-retest reliability?
3. Which section(s) of the Independent Living Skills Assessment Instrument has/have the lowest level of test-retest reliability?
4. Does the Independent Living Skills Assessment Instrument as a whole have a significant level of test-retest reliability?
5. Is there a significant level of agreement among three raters for each section on the Independent Living Skills Assessment Instrument?
6. Which section(s) of the Independent Living Skills Assessment Instrument has/have the highest level of agreement among three raters?
7. Which section(s) of the Independent Living Skills Assessment Instrument has/have the lowest level of agreement among three raters?
8. Does the Independent Living Skills Assessment Instrument as a whole have a significant level of agreement among three raters?

### Null Hypotheses

1. There will be no significant level of test-retest reliability on the Independent Living Skills Assessment Instrument.
2. There will be no significant level of interrater reliability among three raters of the Independent Living Skills Assessment Instrument.

### Participants for the Study

Two types of participants were needed for the study: (a) individuals with permanent physical disabilities who were over the age of 16 to whom the ILSAI was administered; and (b) registered occupational therapists to administer the instrument. The former group was recruited from the population of clients at the Center for Independence of the Disabled (CID). The group of occupational therapists was recruited from CID employees and associates affiliated with the agency. The client participants who were selected for test-retest reliability were recruited at the same time as those for interrater reliability.

A letter was sent to 35 current and former clients of CID (see Appendix B). Of the 35 clients approached, 18 volunteered to participate in the study and 15 were selected (five for test-retest

reliability and ten for interrater reliability). The clients were recruited on a voluntary basis and did not receive any independent living skills training during the data collection phase of the study. Clients were not approached if they were currently receiving or in need of independent living skills services. Different clients were used in each part of the study and the raters did not inform any participant of the identities of any other participants in either part of the study.

A sample of five clients was selected for test-retest reliability. The sample consisted of three females and two males who ranged in age from 37 to 58 years old. The following diagnoses were represented: rheumatoid arthritis, multiple sclerosis, amyotrophic lateral sclerosis, cerebral vascular accident, and spinal cord injury (see Table 1). For interrater reliability, a sample of ten clients was selected from the persons who volunteered for the study. The sample consisted of five females and five males who ranged in age from 16 to 28 years old. The following diagnoses were represented: cerebral palsy (five), spina bifida (two), spinal cord injury (two), and arthrogyposis (one) (see Table 2).

The criteria used for selection of raters for test-retest and interrater reliability were that they must be registered occupational therapists and have a minimum of one year of clinical experience working in a physical dysfunction setting. The three raters selected

Table 1

Demographic Data for Test-Retest Client Sample

Client	Disability	Gender	Age
V	Amyotrophic Lateral Sclerosis	F	43
W	Multiple Sclerosis	M	53
X	Rheumatoid Arthritis	F	58
Y	Cerebral Vascular Accident	F	43
Z	Spinal Cord Injury	M	37

Table 2

Demographic Data for Interrater Reliability Client Sample

Client	Disability	Gender	Age
A	Spina Bifida	M	19
B	Spinal Cord Injury	M	27
C	Spina Bifida	M	20
D	Cerebral Palsy	F	27
E	Cerebral Palsy	F	28
F	Cerebral Palsy	F	16
G	Spinal Cord Injury	M	23
H	Cerebral Palsy	F	24
Ja	Cerebral Palsy	F	23
K	Arthrogyposis	M	21

<sup>a</sup>The letter "I" was not used to denote a subject so that it would not be confused with evaluator #1.

met the criteria for selection. Two were employed by CID and were familiar with the ILSAI; they had also been involved in one or more aspects of the creation and revision of the instrument, as well as the development of the protocol for administration. One of these persons was also a graduate student in the Department of Occupational Therapy at San Jose State University and one of the authors of this study. The third evaluator was a registered occupational therapist who was employed by another human services agency in the San Francisco Bay area.

#### Data Collection Techniques for Test-Retest Reliability

The Independent Living Skills Assessment Instrument was not designed to be used as a comparison between individuals, but rather as a method of gathering baseline data on a person's functional ability and then comparing this information with subsequent assessment following intervention by the therapist. As such it was ideally suited to the test-retest method of determining level of reliability.

One evaluator administered the ILSAI two times to each of the five subjects. Administration of the ILSAI took place in the clients' homes. The occupational therapist administered the instrument in sequence as it was written, using the protocol and guidelines for the

scoring categories to determine how to rate each item (see Appendix A). The instrument was administered in an interview format.

There was a minimum of 24 hours between the two administrations of the instrument. The second administration took place as soon as possible, given the scheduling constraints of the client and rater. The retest time ranged from 2 to 5 days, with a mean time of 2.8 days. Prior to administration of the first test, each client signed a written consent to participate in the study (see Appendix C). All five clients selected for test-retest reliability participated for the duration of the study.

Each instrument form was coded by letter in order to protect the confidentiality of the clients. The researchers and selected members of the staff of CID were the only persons who knew the names of the clients who volunteered for the study. The clients were given minimal information prior to and during the study. A debriefing was carried out following administration of the post-test, when the purpose of the study was explained. A copy of the results was made available to each client when the statistical analysis had been completed.

Most items of the ILSAI were used. Those that were excluded were ones where learning might have occurred in the process of administering the instrument. (The excluded items are marked with an asterisk on the ILSAI in Appendix A).

Statistical analysis of the data included a factor analysis of the items by category (i.e., Functional Mobility, Grooming and Hygiene, Dressing, Feeding/Eating, Health, Functional Communication, Object Manipulation, Homemaking, and Community Involvement). A single factor analysis of variance was used in the statistical analysis of the data.

#### Data Collection Techniques for Interrater Reliability

In the interrater reliability test, three raters each administered the ILSAI one time to the same ten clients. A meeting was held prior to the collection of data with all three occupational therapists present. The purpose of the meeting was to review the instrument and protocol, and to agree on terminology and method of administration. Each therapist received coded instrument forms, a master list with clients' names with letter codes (used for confidentiality), and a table which outlined order of administration of the ILSAI for each client. The order of administration was rotated in order to eliminate the order effect on the results of the study (see Table 3). Each instrument form was coded by letter to protect the confidentiality of the clients. The researchers and selected members of the staff of the Center for Independence of the Disabled were the



Table 3

Order of Administration of ILSAI for Interrater Reliability

Client	Rater		
A	#1	#2	#3
B	#1	#3	#2
C	#1	#2	#3
D	#1	#3	#2
E	#2	#3	#1
F	#2	#1	#3
G <sup>a</sup>	#2	#3	#1
H	#3	#2	#1
J <sup>b</sup>	#3	#1	#2
K	#3	#2	#1

<sup>a</sup>Subject G withdrew himself from the study following the second administration of the ILSAI.

<sup>b</sup>The letter "I" was not used to denote a subject so that it would not be confused with evaluator #1.

only persons who knew the names of the clients who volunteered for the study.

The raters did not discuss any experiences related to the data collection while the ILSAI was being administered. The only contact related to the data collection involved one rater contacting the next rater after she had administered the instrument. Each rater administered the ILSAI in sequence as it was written, using the protocol and guidelines for the scoring categories to determine how to rate each item.

Administration of the ILSAI took place in the clients' homes. The instrument was administered in an interview format. There was a minimum of 24 hours between each administration of the instrument. Each administration took place as soon as possible, given the scheduling constraints of the client and rater. The retest time ranged from 15 to 43 days, with a mean time of 28.9 days.

Prior to the first administration of the ILSAI, each client signed a written consent to participate in the study (see Appendix C). One client withdrew from the study following the second administration of the instrument. The other nine clients participated fully in the study. The clients were given minimal information prior to and during the study. A debriefing was carried out following the third administration of the ILSAI, when the purpose of the study was

explained. A copy of the results was made available to each client when the statistical analysis had been completed.

Most items of the ILSAI were used. Those that were excluded were ones where learning might have occurred in the process of administering the instrument. (The excluded items are marked with an asterisk on the ILSAI in Appendix A). Statistical analysis of the data included a factor analysis of the items by category (i.e., Functional Mobility, Grooming and Hygiene, Dressing, Feeding/Eating, Health, Functional Communication, Object Manipulation, Homemaking, and Community Involvement.) A single factor analysis of variance and the Tukey method of multiple comparisons were used in the statistical analysis of the data.

## CHAPTER 4

### ANALYSIS OF DATA

#### Purpose

The purpose of this analysis was to statistically evaluate the data collected for test-retest reliability and interrater reliability. The following questions were addressed:

1. Is there a significant level of test-retest reliability for each section of the Independent Living Skills Assessment Instrument?
2. Which section(s) of the Independent Living Skills Assessment Instrument has/have the highest level of test-retest reliability?
3. Which section(s) of the Independent Living Skills Assessment Instrument has/have the lowest level of test-retest reliability?
4. Does the Independent Living Skills Assessment Instrument as a whole have a significant level of test-retest reliability?
5. Is there a significant level of agreement among three raters for each section on the Independent Living Skills Assessment Instrument?

6. Which section(s) of the Independent Living Skills Assessment Instrument has/have the highest level of agreement among three raters?

7. Which section(s) of the Independent Living Skills Assessment Instrument has/have the lowest level of agreement among three raters?

8. Does the Independent Living Skills Assessment Instrument as a whole have a significant level of agreement among three raters?

(The questions related to the interrater reliability study were appropriate only if the analysis concluded that there was a high degree of test-retest reliability.)

The null hypotheses were stated as follows:

1. There will be no significant level of test-retest reliability on the Independent Living Skills Assessment Instrument.

2. There will be no significant level of interrater reliability among three raters of the Independent Living Skills Assessment Instrument.

## Analysis Approach

### Test-Retest Reliability

The test-retest reliability part of the study consisted of a single rater administering the Independent Living Skills Assessment Instrument (ILSAI) to five clients on two separate occasions. For each client the statistic of interest was defined as Percentage Discrepancy (PD), where PD is equal to the percentage of questions in each section of the instrument that were answered differently across two applications of the instrument. The PD was computed for each of nine sections of the instrument for all five clients. From the PD values, the sections with the highest and lowest levels of reliability were identified (the smaller the PD, the higher the level of test-retest reliability).

Next, a general single-factor analysis of variance was employed to determine if the difference between the average PD values for each section was significantly different for the sections. Analysis of variance models were used to analyze effects of the independent variable under study on the dependent variable. In this analysis, the independent variable was the instrument section and the dependent variable was the PD. If it was determined that there were significant differences between the mean responses by type of section, the

Tukey method of multiple comparisons was the formula of choice to determine which of the sections was statistically different from the others.

Finally, the computation of the average PD across the instrument and the construction of confidence intervals was used to determine the significance of the instrument as a whole, as well as section by section. The confidence coefficient used was 90% (or, the level of significance,  $\alpha = .10$ ). The interpretation of the confidence coefficient would be that if the experiment were conducted an infinite number of times, the mean PD computed for each experiment would fall in the interval constructed with the confidence coefficient. Thus upper bounds on the PD could be formed for a particular confidence coefficient.

#### Interrater Reliability

The interrater reliability portion of the study consisted of three raters administering the ILSAI to nine clients. In the analysis of these data, the PD was defined to be the percent of questions in each section for which responses did not match across the three administrations of the instrument to the same client. The rest of the analysis was conducted as described above for the test-retest reliability analysis.

### Results for Test-Retest Reliability

The summary of the data collected from the test-retest reliability study has been listed in Table 4. The following information was derived from these data:

$$\text{Mean number of discrepancies} = (2 + 2 + 8 + 0 + 8) / 5 = 4$$

$$\text{Mean Percent Discrepancy (PD)} = 4/123 = 0.0325 \text{ or } 3.25\%$$

$$\text{Standard Error of estimate} = .0304 \text{ or } 3.04\%$$

With the data in Table 4, 90% confidence limits were constructed on the estimated PD using the following formulas:

$$\text{Lower bound of interval} = \text{Mean estimate} - \text{Standard Error} \times t_{.90,4}$$

$$\text{Upper bound of interval} = \text{Mean estimate} + \text{Standard Error} \times t_{.90,4}$$

(where the value of the 't' statistic with 4 degrees of freedom was obtained from standard t distribution tables).

With these formulas, the interval for the mean PD was found to be within the low to high parameters of 0.0% and 9.7% at the 90% level of confidence. This interval applied to the instrument as a whole, ignoring any special effects due to type of section. The



Table 4

Number of Discrepancies for Test-Retest Reliability

Section	Number of Questions	Clients					Totals
		V	W	X	Y	Z	
I	47	1	0	1	0	4	6
II	16	0	0	1	0	1	2
III	17	0	0	2	0	1	3
IV	5	0	0	0	0	0	0
V	1	0	0	0	0	0	0
VI	7	1	2	0	0	0	3
VII	4	0	0	1	0	0	1
VIII	24	0	0	3	0	2	5
IX	2	0	0	0	0	0	0
Totals	123	2	2	8	0	8	20

interpretation of the interval was that if the instrument were administered an infinite number of times, the mean PD (percentage of discrepancies between the test and retest) would be less than 9.7% in 90% of these administrations. In other words, the test-retest reliability would be higher than 90.3% in 90% of the administrations. The same formula was used to determine the predicted reliability of each section at the 90% level of confidence (see Table 5).

To examine the effects of the type of section on the PD in the test-retest reliability, the PD for each section was computed for all clients (see Table 6). The responses to Sections IV, V, and IX contained no discrepancies across the two administrations of the test. Section VI had the highest level of discrepancy (and thus the lowest test-retest reliability) with a PD of 12.8%. To determine if these differences in the mean PD by section were significant at the 90% level of confidence, a single-factor analysis of variance was performed.

The test statistic was

$F^* = MSR/MSE$  where

$MSR = (\sum n_i (\text{Mean PD}_i - \text{Overall Mean PD})^2) / (\text{Number of sections} - 1)$

and

Table 5  
Predicted Test-Retest Reliability by Section

Section	Number of Questions	Reliability <sup>a</sup>
I	47	90.0%
II	16	90.2%
III	17	85.3%
IV	5	100%
V	1	100%
VI	7	64.2%
VII	4	71.3%
VIII	24	83.3%
IX	2	100%

<sup>a</sup>Predicted reliability at the 90% confidence level.

Table 6

Percent Discrepancy (PD) for Test-Retest Reliability

Section	Clients					Total	Mean	S.E. <sup>a</sup>	L <sup>b</sup>	U <sup>c</sup>
	V	W	X	Y	Z					
I	.02	.00	.02	.00	.08	.13	.03	.04	.00	.10
II	.00	.00	.06	.00	.06	.12	.02	.03	.00	.10
III	.00	.00	.12	.00	.06	.18	.04	.05	.00	.15
IV	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
V	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
VI	.14	.29	.00	.00	.00	.43	.09	.13	.00	.36
VII	.00	.00	.25	.00	.00	.25	.05	.11	.00	.29
VIII	.00	.00	.12	.00	.08	.21	.04	.06	.00	.18
IX	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00

<sup>a</sup>Standard Error<sup>b</sup>Lower Bound on 90% confidence interval<sup>c</sup>Upper Bound on 90% confidence interval

$$\text{MSE} = (\sum_i \sum_j (\text{PD}_{ij} - \text{Mean PD}_i)^2 / (\text{Number of computed PDs} - \text{Number of sections}),$$

and 'i' represented 1 up to the number of sections (i.e., 9) and 'j' represented 1 up to the number of clients (i.e., 5). The Mean PD<sub>i</sub> referred to the mean PD for section i. The overall mean PD was the average of all Mean PD<sub>i</sub>. The number of computed PDs was 45 (one PD for each section for each client). From the data in Table 6, the test statistic F\* was computed to be .98. This value was less than 1.88, which indicated that there was no significant difference in the PD due to the type of section.

In summary, 1) The overall percent discrepancy was less than 9.7% at the 90% level of confidence. Therefore, the test-retest reliability will be found to be greater than 90.3% in 90% of the times the instrument is administered; 2) Sections IV, V, and IX had the highest test-retest reliability (100%), while Section VI had the lowest test-retest reliability (64.2%). The differences in the test-retest reliability for the different sections was not found to be significant at the 90% confidence level.

### Results for Interrater Reliability

Working under the assumption that a repeatability of 90.3% was acceptable for a study of this nature, the interrater reliability data were analyzed. In this analysis, PD was defined to be the percentage of questions whose responses did not match across the three applications of the instrument to the same client. The summary of the data for the nine clients is in Table 7. The following information was derived from these data:

$$\text{Mean number of discrepancies} = (4 + 11 + 13 + 35 + 28 + 27 + 31 + 23 + 44)/9 = 24.$$

$$\text{Mean Percent Discrepancy (PD)} = 24/123 = 0.1951 \text{ or } 19.51\%$$

$$\text{Standard Error of estimate} = .1031 \text{ or } 10.31\%$$

The formulas for confidence intervals presented in the earlier section on the test-retest reliability were also used for the analysis of interrater reliability. The interval for the mean PD for the interrater reliability was found to be within the low to high parameters of 0.0% and 38.7% at the 90% level of confidence. (The number of degrees of freedom for the 't' value was 8.) As with the test-retest reliability confidence interval, this interval was applied to the instrument as a whole, ignoring any special effects due to type of section. The interpretation of the interval was that if the instrument were

Table 7

Number of Discrepancies for Interrater Reliability

Section	# of Questions	Clients									
		A	B	C	D	E	F	H	J	K	
I	47	2	5	5	17	4	11	8	5	13	
II	16	0	2	0	3	2	2	3	5	2	
III	17	1	0	4	6	6	3	6	3	5	
IV	5	0	0	0	1	0	0	0	3	3	
V	1	0	0	0	0	0	1	0	0	1	
VI	7	0	0	1	0	1	2	2	1	0	
VII	4	0	0	0	0	1	0	3	1	0	
VIII	24	0	3	1	7	12	6	8	3	18	
IX	2	1	1	2	1	2	2	1	2	2	
Totals	123	4	11	13	35	28	27	31	23	44	

administered an infinite number of times by these three raters, the mean PD (percentage of discrepancies in the responses from the clients) would be less than 38.7% in 90% of these experiments. In other words, the reliability was greater than 61.3% at the 90% level of confidence. The same formula was used to determine the predicted reliability of each section at the 90% level of confidence (see Table 8).

To examine the effects of the type of section on the PD in the interrater reliability, the PD for each section for all clients was computed. Table 9 summarizes these data. Then the mean PD was derived using the same method as in the analysis of test-retest reliability (see Table 10). Section VI had the lowest mean PD at 11.1%. Section IX had the highest percentage of discrepancies at 77.8%. To determine if these differences in the mean were significant at the 90% level of confidence, a single factor analysis of variance was performed.

The test statistic was

$F^* = MSR/MSE$  where

$MSR = (\sum n_i (\text{Mean PD}_i - \text{Overall Mean PD})^2 / (\text{Number of sections} - 1)$

and



Table 8  
Predicted Interrater Reliability by Section

Section	Number of Questions	Reliability <sup>a</sup>
I	47	64.0%
II	16	69.0%
III	17	53.4%
IV	5	36.0%
V	1	0.0%
VI	7	66.8%
VII	4	39.0%
VIII	24	28.8%
IX	2	0.0%

<sup>a</sup>Predicted reliability at the 90% confidence level.

Table 9

Percent Discrepancy (PD) for Interrater Reliability for Each Client

Section	N <sup>a</sup>	Clients								
		A	B	C	D	E	F	H	J	K
I	47	.04	.11	.11	.36	.08	.23	.17	.11	.28
II	16	.00	.12	.00	.19	.12	.12	.19	.31	.12
III	17	.06	.00	.24	.35	.35	.18	.35	.18	.29
IV	5	.00	.00	.00	.20	.00	.00	.00	.60	.60
V	1	.00	.00	.00	.00	.00	1.00	.00	.00	1.0
VI	7	.00	.00	.14	.00	.14	.28	.28	.14	.00
VII	4	.00	.00	.00	.00	.25	.00	.75	.25	.00
VIII	24	.00	.12	.04	.29	.50	.25	.33	.12	.75
IX	2	.50	.50	1.00	.50	1.00	1.00	.50	1.00	1.00

<sup>a</sup>N equals the number of items in each section.

Table 10

Mean Percent Discrepancy (PD) for Interrater Reliability

Section	# of Questions	Total	Mean	S.E. <sup>a</sup>	L <sup>b</sup>	U <sup>c</sup>
I	47	1.49	0.16	0.10	0.00	0.36
II	16	1.19	0.13	0.10	0.00	0.31
III	17	2.00	0.22	0.13	0.00	0.46
IV	5	1.40	0.16	0.27	0.00	0.64
V	1	2.00	0.22	0.45	0.00	1.00
VI	7	1.00	0.11	0.12	0.00	0.33
VII	4	1.25	0.14	0.25	0.00	0.61
VIII	24	2.42	0.27	0.25	0.00	0.71
IX	2	7.00	0.78	0.26	0.29	1.00

<sup>a</sup>Standard Error<sup>b</sup>Lower Bound on 90% confidence level<sup>c</sup>Upper Bound on 90% confidence level

$$\text{MSE} = (\sum_i \sum_j (\text{PD}_{ij} - \text{Mean PD}_i)^2 / (\text{Number of computed PDs} - \text{Number of sections}),$$

and 'i' represented 1 up to the number of sections (i.e., 9) and 'j' represented 1 up to the number of clients (i.e., 5). The Mean PD<sub>i</sub> referred to the mean PD for section i. The overall mean PD was the average of all Mean PD<sub>i</sub>. The number of computed PDs was 81 (one PD for each section for each client).

From the data in Table 10, the test statistic F\* was computed to be 6.86, which was greater than 1.77, which indicated that there were significant effects on response for the type of section. To determine whether the mean percentage discrepancy for each section was significantly different from each other section, the Tukey method of multiple comparisons was used. The method involved setting up confidence intervals for the difference in means between all possible pairs of sections. If the confidence interval thus formed contained zero, the difference between the sections was not significant; if not, it was significant. Since there were nine sections, 36 possible pairs were examined.

The Tukey method utilizes the studentized range distribution. (This distribution may be found in standard statistical tables.) The interval was formed around the difference in the mean PDs of a pair

of sections. The interval was computed as follows (comparing section i and section j):

$$\text{Lower Bound} = (\text{Mean PD}_i - \text{Mean PD}_j) - \text{S.E.}(i,j) * T$$

$$\text{Upper Bound} = (\text{Mean PD}_i - \text{Mean PD}_j) + \text{S.E.}(i,j) * T$$

where

S.E.(i,j) was the standard error of the difference in the means and was equal to  $\sqrt{\text{MSE}/9}$ ,

and

T was the Tukey statistic equal to  $1/\sqrt{2} * q(90\%, 9, 72)$ . The 'q' refers to the studentized range distribution and the three parameters are the confidence level, the number of sections, and the number of computed PDs minus the number of sections.

Table 11 shows the results of the formation of these intervals. As the table data illustrate, the mean PD of sections I and IV, sections II and VII, and sections IV and VII were not significantly different from each other.

In summary, 1) The overall percent discrepancy was less than 38.7% at the 90% level of confidence, therefore the interrater reliability would be greater than 61.3% in 90% of the cases when the instrument was administered; 2) Section II demonstrated the highest predicted reliability at 69.0%, while section IX had the lowest predicted reliability of 0%; 3) There was a significant difference for

Table 11

Pairwise Comparisons of Sections Using the Tukey Method

Sections	Lower	Upper	N <sup>a</sup>	Sections	Lower	Upper	N
I, II	0.02	0.05		III, VII	0.06	0.10	
I, III	-0.08	-0.04		III, VIII	-0.06	-0.03	
I, IV	-0.01	0.03	N	III, IX	-0.57	-0.54	
I, V	-0.08	-0.04		IV, V	-0.08	-0.05	
I, VI	0.04	0.07		IV, VI	0.03	0.06	
I, VII	0.01	0.04		IV, VII	0.00	0.04	N
I, VIII	-0.12	-0.08		IV, VIII	-0.13	-0.10	
I, IX	-0.63	-0.59		IV, IX	-0.64	-0.60	
II, III	-0.11	-0.07		V, VI	0.09	0.13	
II, IV	-0.04	0.00		V, VII	0.06	0.10	
II, V	-0.11	-0.07		V, VIII	-0.06	-0.03	
II, VI	0.00	0.04		V, IX	-0.57	-0.54	
II, VII	-0.02	0.01	N	VI, VII	-0.05	-0.01	
II, VIII	-0.16	-0.12		VI, VIII	-0.18	-0.14	
II, IX	-0.66	-0.63		VI, IX	-0.68	-0.65	
III, IV	0.05	0.08		VII, VIII	-0.15	-0.11	
III, V	-0.11	-0.07		VII, IX	-0.66	-0.62	
III, VI	0.09	0.13		VIII, IX	-0.53	-0.49	

<sup>a</sup>Signifies an absence of any statistically significant difference between the two section mean PDs

the type of section in the reliability of the responses in all except between sections I and IV, sections II and VII, and sections IV and VII.

## CHAPTER 5

### RESULTS, IMPLICATIONS, SUMMARY AND RECOMMENDATIONS

#### Discussion

In this discussion, the eight questions posed in this study of test-retest and interrater reliability are discussed in detail. Also included in this discussion are problems and concerns the authors found with the Independent Living Skills Assessment Instrument (ILSAI) and its administration while carrying out the study, as well as the possible effects on the outcome of the study. Finally, studies included in the literature review are discussed in relation to this study.

There were eight questions and two hypotheses posed for this study. For the purposes of this study the authors considered any score greater than 85% to be an acceptable level of reliability for both test-retest and interrater reliability. The following are the answers to the questions:

1. There was a statistically significant predicted test-retest reliability of the Independent Living Skills Assessment Instrument



for sections I (90.0%), II (90.2%), III (85.3%), IV (100%), V (100%), and IX (100%).

2. Sections IV, V, and IX of the Independent Living Skills Assessment Instrument showed the highest level of test-retest reliability at 100%.

3. Section VI of the Independent Living Skills Assessment Instrument had the lowest test-retest reliability at 64.2%.

4. Overall, the Independent Living Skills Assessment Instrument had a statistically significant test-retest reliability of 90.2%.

5. None of the individual sections of the Independent Living Skills Assessment Instrument had statistically significant levels of interrater reliability.

6. Sections II and VI of the Independent Living Skills Assessment Instrument had the highest level of interrater reliability at 69.0% and 66.8% respectively.

7. Sections V and IX of the Independent Living Skills Assessment Instrument had the lowest level of interrater reliability at 0%.

8. The Independent Living Skills Assessment Instrument had an overall predicted interrater reliability of 61.3% which was not statistically significant.

The following are the results of hypothesis testing:

1. Significance levels for test-retest reliability for some categories and for the Independent Living Skills Assessment

Instrument as a whole were reached; therefore, Hypothesis 1 was refuted.

2. Significance levels were not reached in interrater reliability among three raters of the Independent Living Skills Assessment Instrument; therefore, Hypothesis 2 was accepted.

In the test-retest portion of the study, Sections VI, VII, and VIII had a predicted reliability of less than 85%, and therefore were not considered to be reliable. It is not known why these three sections fell below the acceptable level while the tool as a whole had an acceptable level of reliability. One factor that may have affected the results was the use of a small sample for data collection. The small sample size may have had a different effect on the outcome of each section, as opposed to the outcome for all the sections combined.

It is interesting to note, however, that after performing the single factor analysis of variance, it was determined that the difference in the test-retest reliability between the different sections was not significant. So although these three sections fell below the acceptable level of test-retest reliability, the differences between these three sections and the other six sections, which were considered reliable, were not significant.

The three sections with the highest percentages of test-retest agreement were also among the four categories with the fewest items per category (in part due to items being omitted because of the

factor of learning). The number of items in these three sections ranged from one to seven (as opposed to 49 items in Section I). Section VI, which had the lowest level of test-retest reliability, had only seven items. The small number of items may have affected the very high and low sectional outcomes of test-retest reliability.

For interrater reliability, the two sections with the lowest level of reliability are the two sections with the fewest items; Section V had one item and Section IX had two items. With fewer items a chance error has less probability of being balanced with another chance error in the opposite direction and can affect the overall measurement. The number of items in these categories may have affected the outcome of the predicted interrater reliability. Section II, which measures grooming and hygiene, had 16 items and had the highest reliability at 69.0%. This result was well above that of the lowest score for a section, but was still an unacceptable level of reliability for the purposes of this study.

The outcomes of this study would appear to mean, in practical terms to occupational therapists, that the ILSAI may be reliable in its entirety with one rater performing both the pretest and the posttest. The ILSAI cannot be considered a reliable instrument for comparison purposes beyond measuring a client's performance in posttest against his/her performance in pretest.

Problems and concerns with the instrument and its administration were studied for their possible effects on the outcome of the study, especially regarding the interrater reliability outcome. First, the protocol for the ILSAI gives no instruction as to the method of administration of the instrument (i.e., interview or observation), only the description of an acceptable behavior for each item. The authors eliminated this confusion for the study by agreeing to instruct the participating raters to administer the instrument by interview only. Thus the results should not have been affected by that lack of direction.

The directions for scoring, however, may have been confusing for the raters as well as the interpretation of the scoring categories. Three categories in particular appear to be unclear and may overlap in their interpretation and scoring between raters. These categories are "Able To But Doesn't Choose To," "Don't Know," and "Not Applicable." An example of a different interpretation could be the scoring of a client who says he knows how to ride the bus, but a relative drives him on all of his errands. This may evoke a response of "Able To But Doesn't Choose To" from one rater, while it could be scored as "Not Applicable" by another rater. These scoring categories may need to be eliminated or combined into one larger category to prevent this ambiguity.

Another question raised in the minds of the authors was whether or not the level of familiarity with the assessment instrument would affect the outcome of the study. The assumption was that because Rater #1 and Rater #3 were very familiar with the tool, they would agree more often when scoring items. In fact this was not the case. When further analysis of the results was performed it was found that of 216 occasions when the three raters did not agree, 10.6% of the time all three raters selected different responses. Raters #1 and #2 agreed 26.9% of the time, Raters #1 and #3 agreed 22.2% of the time, and Raters #2 and #3 agreed 40.3% of the time. The results suggest that the raters' level of experience with the instrument was not a factor. They also indicate that at least two of the raters agreed in most cases. It is not known why Rater #2 and #3 had such a low level of agreement when compared with Rater #1's agreement with Raters #2 and #3.

During this analysis it was also noted that in 33.8% of the cases where one rater did not agree with the other two, the response "Not Applicable" was one of the categories selected when the item was scored. As discussed before, such a high percentage of discrepancy could suggest that having this particular response available may have complicated the interrater reliability outcomes.

Last, a look was taken at the time intervals between administrations of the ILSAI for both test-retest and interrater

reliability. In the literature review, McLaughlin (1980) noted that the issue of establishing an appropriate interval between administrations of the instrument must be addressed. The authors used a minimum of 24 hours between each administration of the ILSAI for both parts of the study. The second administration was given as soon as possible after that time period because the authors wanted to minimize the possibility of a change in functional status.

The retest time for the test-retest reliability ranged from 2-5 days, a relatively short span of time. The time between raters' administrations of the ILSAI for interrater reliability, however, ranged from 15-43 days due to scheduling constraints. It should be noted that this is a much longer time period between administrations of the instrument than for the retest administrations in the test-retest portion of the study. This longer interval may have affected the interrater reliability outcomes.

It was not possible to compare the outcome of this reliability study with studies on comparable instruments because those studies used different statistical analyses to measure the reliability outcomes. One study, however, that of Watts, Keilhofner, Bauer, Gregory, and Valentine (1986), included both test-retest and interrater reliability testing. The interrater part of the reliability study included a pilot study using 13 clients and two occupational therapists as raters. This is much closer to the size of the ILSAI

interrater reliability study (with 10 clients and 3 occupational therapists), than their second larger and final interrater reliability study (with 83 clients). The researchers in this study view this investigation of the ILSAI as similar to that of a pilot study, and as a point of departure for improvement and further development of the ILSAI.

Most of the other studies cited in the literature review had test-retest or interrater reliability outcomes that were acceptable to the researchers. For those that were not, little discussion was found on possible methods to improve reliability for the specific instruments or plans for further development of the instrument.

The authors of this study wish to emphasize the preliminary nature of this research, which is considered only a step in the total development of the ILSAI. Because the evaluation instrument had not been studied for acceptable reliability and because of the limited resources of the Center for Independence of the Disabled (CID), this study was designed to provide a foundation of information and feedback from which development of the instrument could continue as well as to give information for its current use. The authors also wish to recognize that the results of the study are limited to the items tested. Some items of the ILSAI were eliminated because the factor of learning could have affected the responses.

### Implications for the Profession

Occupational therapy practice is based on the belief that the functioning of a disabled individual can improve given specific purposeful activity chosen by the occupational therapist in conjunction with the client following accurate evaluation. Critical to this process is the administration of evaluation instruments and the accuracy and reliability of the evaluation tool itself.

Many evaluation instruments, including the ILSAI, are used as a pretest to determine the initial level of functioning and as a post-test to determine the progress of the client's treatment. Without a reliable evaluation tool, the quality of care, the direction of treatment, and the client's progress will be negatively affected. With an evaluation tool that has been developed and tested for standardization, including reliability, the occupational therapist can evaluate the client's functioning, treat the client in specific areas needed, reevaluate the client's functioning, and develop a plan for further treatment, if necessary. In addition, with reliable evaluation instruments, occupational therapists can share information about a client with other occupational therapists and health professionals with the understanding that the occupational therapist is reporting from a common, accurate baseline of information.



In addition, the reality of responsibility for funding occupational therapy programs is becoming more and more a direct responsibility of the occupational therapist. Even though the occupational therapist may not be directly involved with administration of the department, the occupational therapist plays a vital role in the direction of the profession by keeping accurate records of clients' status and outcomes of treatment. In most programs, funding measures result directly from the occupational therapist's ability to evaluate and record the client's function from admission through discharge.

Reliable evaluation tools are the cornerstone of this process. Occupational therapy can only continue to grow if it fills the need for reliable assessment instruments specific to each area of practice.

The ILSAI can be of particular interest to occupational therapists working in the community. There appears to be a gap in services between hospital and community and/or sheltered living and employment and full functioning in the community. The ILSAI is specifically designed for use with clients who are in that position. This particular evaluation instrument was designed specifically for the population who are now living in the community but need improvement of functioning in one or more areas of their life. This tool needs to be reliable for both the occupational therapists who are treating these clients for the funding of this type of service and for

the clients themselves, who need improvement in their ability to live as independent members of the community.

This preliminary study of reliability of the ILSAI has established its test-retest reliability for a limited population. Occupational therapists can now use this evaluation, knowing some of its limitations and usefulness as an evaluation tool in a specific setting. Hopefully, this preliminary study will also demonstrate the importance of careful and deliberate attempts to establish and report reliability of occupational therapy assessment tools and encourage occupational therapists to further develop this and other instruments with specific goals for accurate measurement.

### Summary

A test-retest and interrater reliability study was carried out on the Independent Living Skills Assessment Instrument, a functional capacity evaluation which is used at the Center for Independence of the Disabled in Belmont, California. The ILSAI measures a comprehensive range of functional living skills and was administered by registered occupational therapists. The other participants in the study were persons with permanent physical disabilities who were living in the community. These participants were volunteers who were recruited from the client files at CID.

The test-retest reliability consisted of two administrations of the instrument by one rater to each of five subjects. The interrater reliability portion of the study involved three subjects, all registered occupational therapists, who administered the instrument one time to each of ten participants.

Overall test-retest reliability of the instrument was 90.2% at the 90% level of confidence, which is greater than the acceptable reliability level of 85%. The nine sections of the instrument had predicted reliabilities ranging from 64.2% to 100% at the 90% confidence level. Three sections of the instrument fell below the acceptable level of test-retest reliability.

Overall interrater reliability of the instrument was found to be 61.3% at the 90% confidence level. This score, as well as predicted reliability for each of the nine sections, which ranged from 0% to 69%, did not meet the minimum acceptable level of reliability.

Problems and concerns with the ILSAI and its administration found by the authors during this study include possible confusion in the interpretation of the protocol and directions for administering the instrument, possible confusion over interpretation of several of the scoring categories and the small number of items in certain sections of the evaluation.

The authors considered this study to be preliminary testing of test-retest and interrater reliability of the ILSAI and recommend

that the instrument not be used for formal evaluation of clients by more than one rater until it is revised and retested for acceptable interrater reliability.

### Recommendations

It is recommended that the protocol and instructions for the ILSAI be revised to include a guide for administration. This guide would include the information that the researchers gave verbally to the raters and clearer, more specific definitions of each of the items to eliminate variation in the rater's response. It is also recommended that consideration be given to eliminating one or more of the following categories for scoring: 1) "Able To But Doesn't Choose To," 2) "Don't Know," 3) "Not Applicable," or that these categories be defined more specifically so that their interpretation can be more precise. Consideration should also be given to revising the content of the instrument by adding more items of equal kind and quality. This would be especially important in those sections of the ILSAI where there are very few items.

Although the original intent that motivated the exclusion of certain items from testing in the study was to avoid a confounding variable of learning, it may be important to test the entire

instrument for reliability. It is therefore recommended that the eliminated items be restored for future reliability testing.

After the revisions in the protocol and the ILSAI are made, it is recommended that both test-retest and interrater reliability testing be conducted using a larger population of clients. If possible, random sampling should be employed so that the results can be generalized to a wider population.

If the ILSAI is used before revisions and retesting can be completed, its limitations must be recognized. In such cases, it is recommended that the instrument be used in its entirety rather than by section and pre- and post-test administration should be performed by the same occupational therapist. If more than one therapist is treating a client, either simultaneously or in succession, each therapist should complete his/her own pre- and post-tests to record the client's baseline functioning and the outcome of intervention.

## References

- American Occupational Therapy Association. (1981). Entry-level role delineation for OTRs and COTAs. Rockville, MD: Author.
- American Occupational Therapy Association. (1972). Occupational therapy: Its definitions and functions. Rockville, MD: Author.
- Anastasi, A. (1976). Psychological testing (4th ed.). New York: Macmillan.
- Baum, C. M. (1980). Independent Living: A critical role for occupational therapy. The American Journal of Occupational Therapy, 34, 773-774.
- Brorsson, B., & Asberg, K. H. (1984). Katz index of independence in ADL: Reliability and validity in short-term care. Scandinavian Journal of Rehabilitation Medicine, 16, 125-132.
- Center for Independence of the Disabled. (1983). Mission statement of the Center for Independence of the Disabled, Inc. Belmont, CA: Author.
- Charlton, J. R., Patrick, D. L., & Peach, H. (1983). Use of multivariate measures of disability in health surveys. Journal of Epidemiology and Community Health, 37, 296-304.
- DeJong, G. (1979). Independent living: From social movement to analytic paradigm. Archives of Physical Medicine and Rehabilitation, 60, 435-446.

- Jette, A. M. (1980). Functional status index: Reliability of a chronic disease evaluation instrument. Archives of Physical Medicine and Rehabilitation, 61, 395-401.
- Keith, R. A. (1984). Functional assessment measures in medical rehabilitation: Current status. Archives of Physical Medicine and Rehabilitation, 65, 74-78.
- Klein, R. M., & Bell, B. (1982). Self care skills: Behavioral measurements with Klein-Bell ADL Scale. Archives of Physical Medicine and Rehabilitation, 63, 335-338.
- Kuriansky, J., & Gurland, B. (1982). The performance test of activities of daily living. International Journal of Aging and Human Development, 7, 343-352.
- McLaughlin, J. (1981). Self-care behavior of persons with multiple sclerosis in Denmark: Development of a behavioral instrument (Doctoral dissertation, Southern Illinois University at Carbondale, 1981). Dissertation Abstracts International, 42, 1583A.
- Neter, J., Wasserman, W., & Kutner, M. H. (1985). Applied Linear Statistical Models (2nd ed.). Homewood, IL: R. D. Irwin.
- Peterson, C., Shriver, D., & Farace, J. (1981). Occupational therapy's role in independent or alternate living situations (Official Position Paper). Rockville, MD: American Occupational Therapy Association.

- Reed, K. L., & Sanderson, S. R. (1980). Concepts of Occupational Therapy. Baltimore, MD: Williams and Wilkins.
- Stein, F. (1980). Anatomy of Research in Allied Health. Cambridge, MA: Schenkman.
- Sparrow, S. S., & Cicchetti, D. V. (1978). Behavior rating inventory for moderately, severely, and profoundly retarded persons. American Journal of Mental Deficiency, 82, 365-374.
- Verville, R. E. (1979). Federal legislative history of independent living programs. Archives of Physical Medicine and Rehabilitation, 60, 447-451.
- Watts, J. H., Kielhofner, G., Bauer, D. F., Gregory, M. D., & Valentine, D. B. (1986). The assessment of occupational functioning: A screening tool for use in long-term care. American Journal of Occupational Therapy, 40, 231-240.
- Webster's New Collegiate Dictionary. (1977). Springfield, MA: G. & C. Merriam.
- Wolber, G., & Lira, F. T. (1981). Relationship between Bender Designs and basic living skills of geriatric psychiatric patients. Perceptual and Motor Skills, 52, 16-18.



Appendix A

Independent Living Skills Assessment Instrument,  
Protocol, and Instructions

The items that are marked with an asterisk (\*) were not used in this research project because they were identified as items which may have been affected by learning.

## Cover Sheet for Independent Living Skills Assessment Instrument

The purpose of this instrument is to measure the functional independence of the client. All items should be evaluated based on the individual's every day environment and according to their needs at the time of the assessment. This may include, but is not limited to: home, job, and school.

The comments section should be used to note any information which would clarify the assessment (e.g., special method used or equipment used; type of personal assistance needed, etc.). Use of equipment refers to specialized adaptive devices not normally used to perform the activity (e.g., an electric toothbrush is not noted, but a built-up eating utensil is noted).

The following are the criteria which should be used to determine which category to check:

**INDEPENDENT** - Accomplishes task safely and securely without physical assist or supervision.

**PHYSICALLY ABLE, NEEDS SUPERVISION** - Supervision is needed for safety and/or due to cognitive deficits; verbal or observation, but no physical assistance

**NEEDS PERSONAL ASSIST** - Physical assist needed to accomplish task

A - Able to direct assistant

U - Unable to direct assistant

**DON'T KNOW** - Ability to accomplish task unknown by evaluator

**ABLE TO BUT DOESN'T CHOOSE TO** - Personal choice to utilize assist, but not necessary

**NOT APPLICABLE** - Task not possible ever, or task not done and not needed or desired

(Cover Sheet, p.2)

The following is a list of the ranking of each category which is graded as a level of independence:

1. Independent  
Needs Personal Assist - A (unfeasible or doesn't want to change)  
Able To But Doesn't Choose To
2. Physically Able, Needs Supervision (unfeasible or doesn't want to)
3. Physically Able, Needs Supervision (feasible to change)
4. Needs Personal Assist - A (feasible to change)
5. Needs Personal Assist - U (unfeasible or doesn't want to change)
6. Needs Personal Assist - U (feasible to change)

(NOTE: Rankings are only used when documenting change in the individual's level of independence (i.e., in a pre-test/post-test situation). Change in level of independence is recorded with "+" (positive change), "0" (no change), or "-" (negative change). Judgment is not made regarding the degree of change.)

## INDEPENDENT LIVING SKILLS ASSESSMENT INSTRUMENT PROTOCOL

## I. FUNCTIONAL MOBILITY

## A. Transfers and Positioning

## 1. Bed

- a. into - can move from sitting or standing position to sitting or lying on bed
- b. out of - reverse process
- c. positioning - able to position and reposition self in order to maintain comfort and prevent pressure sores
- d. rolling - able to roll from back to front; reverse process
- e. lie to sit - able to move from lying position to sitting up on bed; reverse process

## 2. Toilet/Commode

- a. on - able to move from sitting or standing position to seated position on toilet
- b. off - reverse process

## 3. Bath/Shower

- a. into - able to move from standing or sitting position into shower or bath; includes getting into position for washing
- b. out of - reverse process

## 4. Chair

- a. on - able to move from standing or sitting position onto any type of chair normally used
- b. off - reverse process
- c. positioning - (refers to wheelchair or regular chair) able to position and reposition self in chair in order to maintain comfort and prevent pressure sores
- d. floor to chair - able to move from any position on floor to sitting position in chair
- e. chair to floor - reverse process

## 5. Car/Van

- a. into - able to move self from standing or wheelchair into car or van, or use wheelchair lift; transfer into seat of vehicle for driving and/or riding, or secure chair in driving and/or riding position
  - b. out of - reverse process
- B. Ambulation (includes walking and crawling)
- 1. Inside - able to ambulate on typical level surfaces, such as rug, linoleum, hardwood floor, carpet, etc.
  - 2. Stairs - able to ambulate up and down stairs
  - 3. Thresholds - able to ambulate over typical thresholds (1" or less)
  - 4. Inclines - able to ambulate up and down standard incline (approximately 12:1 ratio grade), including ramps and curb cuts
  - 5. Escalators - able to get on and off escalator; ride up and down while maintaining balance
  - 6. Elevators - able to enter and exit; press appropriate button(s)
  - 7. Outside - able to ambulate on typical level surfaces, such as concrete, asphalt, etc.; includes moving on sidewalks and pathways
  - 8. Uneven Surfaces - able to ambulate over uneven surfaces, such as grass, gravel, dirt, etc.
  - 9. Curbs - able to ambulate up and down 4-6" curbs
- C. Manual Wheelchair
- 1. Self Propel
    - a. Inside - able to maneuver over typical level surfaces, such as rugs, linoleum, hardwood floor, carpet, etc.
    - b. outside - able to maneuver over typical level surfaces, such as concrete, asphalt, etc.
  - 2. Corners - able to maneuver around any type of corner, inside or outside
  - 3. Thresholds - able to maneuver over typical threshold (1" or less)

4. Inclines - able to maneuver up and down standard inclines (approximately 12:1 ratio grade), including ramps and curb cuts
  5. Elevators - able to enter and exit; press appropriate button(s)
  6. Uneven Surfaces - able to maneuver over uneven surfaces, such as grass, gravel, dirt, etc.
  7. Curbs - able to maneuver up and down 4-6" curbs (i.e., do a "wheelie")
  8. Apply Brakes - able to engage and disengage brakes
  9. Arm Rests - able to unhook fastening mechanism and lift off of and away from chair; reverse process
  10. Leg Rests - able to unhook fastening mechanism and lift leg rest off of and away from chair; reverse process
  11. Load into Vehicle - able to fold or dismantle chair and lift into vehicle or secure onto vehicle
  - \*12. Maintenance - able to perform standard wheelchair maintenance or knows where to go or whom to call
- D. Power Wheelchair (person must also be evaluated for manual wheelchair use)
1. Self Propel
    - a. inside - able to maneuver over typical level surfaces, such as rugs, linoleum, hardwood floor, carpet, etc.
    - b. outside - able to maneuver over typical level surfaces, such as concrete, asphalt, etc.
  2. Corners - able to maneuver around any type of corner, inside or outside
  3. Reverse - able to move backwards in safe, controlled manner
  4. Inclines - able to maneuver up and down standard incline (approximately 12:1 ratio grade)
  5. Elevators - able to enter and exit; press appropriate button(s)
  6. Apply Brakes - able to engage and disengage brakes
  7. Arm Rests - able to unhook fastening mechanism and lift off of and away from chair; reverse process

8. Leg Rests - able to unhook fastening mechanism and lift off of and away from chair; reverse process
9. Load into Vehicle - able to fold or dismantle chair; lift into vehicle or secure in place
- \*10. Maintenance - able to perform standard wheelchair maintenance, including battery care or knows where to go or whom to call

## II. GROOMING AND HYGIENE

- A. Reach Faucet - able to position self at sink and turn water on and off (note if bathroom sink is not used - i.e., not accessible)
- B. Wash Face - able to reach all areas of face; apply sufficient pressure to cleanse skin
- C. Dental Hygiene - able to hold toothbrush; apply toothpaste; reach all regions of mouth; apply sufficient pressure to clean teeth; or able to remove upper and/or lower dentures; prepare solution for soaking; or able to soak dentures
- D. Toileting - able to manipulate clothing before and after using toilet; flush toilet; or able to gain access to catheter and equipment (if applicable)
- E. Toilet Hygiene - able to clean perineal area adequately or able to perform standard maintenance of leg bag and catheter; or able to manage method used
- F. Feminine Hygiene - able to apply and remove sanitary tampons or pads; dispose of soiled tampons or pads
- G. Use Deodorant - able to reach underarm areas; apply deodorant
- H. Shaving/Hair Removal - able to hold razor/shaver; apply shaving cream or soap, if necessary); reach all areas to be shaved; rinse shaved area, rinse or clean razor/shaver; or able to apply and remove depilatory
- I. Use Make-up - able to manipulate various containers and applicators in order to apply desired make-up
- J. Nail Care - able to trim or file fingernails and toenails; apply and remove polish (if used)



K. Bathing

1. Wash - able to reach all areas of body; apply soap; apply sufficient pressure to cleanse skin
2. Dry - able to reach all areas of body with towel; apply sufficient pressure to dry skin

L. Hair

1. Comb/Brush - able to manipulate comb or brush and reach all necessary areas; able to draw through hair
2. Wash - able to wet hair; apply shampoo; massage scalp; rinse completely
3. Dry - able to air dry or towel dry or use dryer
4. Style - able to fix hair in manner desired (may include using blow dryer to style, curling iron, rollers, clips, barrettes, etc.)

III. DRESSING (able to manipulate all types of clothing usually worn)

A. Undergarments

1. Upper

- a. on - able to put on items, including bra, slip, camisole, undershirt, etc.
- b. off - reverse process

2. Lower

- a. on - able to put on items, including underpants, half slip, shorts etc.
- b. off - reverse process

B. Upper Body

1. On - able to put on items, including dress, shirt (button-down, pullover), sweater, vest, jacket, coat, tie, etc.]
2. Off - reverse process

C. Lower Body

1. On - able to put on items, including trousers, pants, shorts, skirt, etc.
2. Off - reverse process

D. Fastening Devices (able to fasten and unfasten devices found on clothing normally worn)

1. Buttons

2. Zippers (e.g., back, jacket, trousers, skirt, etc.)

3. Snaps
  4. Hooks/Eyes
  5. Buckles
  6. Velcro
- E. Shoes - able to put on and take off shoes, including loafers, oxfords, sneakers, boots, etc.
- F. Shoe Fastenings - able to fasten and unfasten, including laces, buckles, zippers, velcro, etc.
- G. Socks/Stockings - able to put on and take off, including socks, panty hose, stockings, knee-his, etc.
- IV. FEEDING/EATING (if client uses feeder, note "I, uses equipment")
- A. Fingers - able to lift food from plate to mouth
- B. Utensils
1. Fork - able to get food onto fork and lift food from plate to mouth
  2. Spoon - able to get food onto spoon (liquid and solid) and lift from dish to mouth
  3. Knife - able to cut food (including meat), spread butter, etc.
- C. Cup/Glass - able to hold cup or glass when filled with liquid; raise to mouth and drink (note "I, uses equipment" if straw is used)
- V. HEALTH
- \*A. Knows Doctor - able to name doctor; knows phone number or where to look it up
- \*B. Knows Dentist - Able to name dentist; knows phone number or where to look it up
- C. Medication (includes prescription/aspirin/vitamins, etc.)
1. Physical Ability - able to open and close containers; handle meds (includes managing liquid to swallow, if applicable)
  - \*2. Manage Schedule - knowledge of when to take meds and proper dosage(s)

- \*D. Skin Care - has knowledge about and ability to check skin in pressure-sensitive areas (note "I, uses equipment" if mirror is used)
- \*E. R.O.M. - has knowledge about and ability to complete range of motion exercises
- \*F. Nutrition Knowledge - able to name the four basic food groups (i.e., dairy, meat and fish, breads and grains, fruits and vegetables)
- \*G. Simple First Aid
  1. Minor Burns - able to answer the following question:  
"What do you do if you burn your hand on a hot pan?"  
(A: Cool with water or apply ice to area.)
  2. Minor Cuts - able to answer the following question:  
"What do you do if you cut your finger with a knife and it starts to bleed?" (A: Apply pressure to stop the bleeding; clean cut by running it under water or using a medicinal cleanser; cover with band aid.)
  3. Dizziness - able to answer the following question: "What should a person do first if he/she starts to feel dizzy while standing?" (A: Sit down or lie down or sit down and put head between legs.)

## VI. FUNCTIONAL COMMUNICATION

- A. Writing - able to write in order to meet functional communication needs (e.g., for homemaking, school, job, correspondence)
- B. Signature - able to complete legible legal signature (note NPA-A if client understands requirement of having witnesses when using identifying mark)
- C. Typing - able to use as means of communication; able to insert and remove paper; depress keys (speed and technique are not factors)
- D. Speaking - able to be understood; includes sign language (if applicable); note NPA if interpreter is used
- E. Turn Pages - able to manipulate pages of reading material in order to turn to next page (books, magazines, newspapers, etc.)

**F. Phone Use****1. Physical Ability**

- a. private - able to pick up receiver; dial or press buttons; hold receiver to ear; put down after use
- b. public - (assume accessible) able to pick up receiver; put in money; dial or press buttons; hold receiver to ear; put down after use

**\*2. Knowledge**

- a. private - able to make call; able to answer phone appropriately
- b. public - knows correct amount of money to put in; knows how to make collect or credit card call

**VII. OBJECT MANIPULATION**

- A. Keys - able to hold key; insert into lock/ignition; turn; remove from lock/ignition (e.g., home, car, etc.)
- B. Money - able to get money from purse or wallet; handle bills and coins; receive money (e.g., change) and put away
- C. Doorknob - able to manipulate any type normally used; able to grasp; turn; pull to open and close door; release
- D. Scissors - able to pick up; hold in position for cutting; cut item (e.g., paper, cloth, etc.); release scissors

**VIII. HOMEMAKING****A. Household**

1. Sweep - able to hold broom; pull across area to be cleaned; manipulate broom and dustpan to pick up debris
- \*2. Vacuum - able to hold and manipulate upright vacuum or nozzle of canister model; push/pull across area to be cleaned; empty and replace bag
3. Dust - able to hold cloth or duster; reach all areas to be dusted; exert sufficient pressure to remove dust
4. Make Bed - able to pull up and adjust covers
5. Change Bed - able to lay out sheets and/or blankets on bed; adjust evenly; tuck in ends and corners; put pillowcase on pillow; place pillow on bed

\*6. Laundry

a. washer - able to insert and remove clothes; add soap or other cleansing agents; set dials; put in money (if coin-operated); start machine (note if machine is at home or laundromat)

b. dryer - (same as washer, except no cleansing agents)

\*7. Iron - able to set up ironing board; understand and adjust controls on iron; add water, if needed; manipulate clothing on board; manipulate iron in order to reach all areas to be pressed

8. Take Out Garbage - able to move garbage to appropriate container outside the home

9. Heavy Cleaning - able to do the following types of cleaning: floors, windows, bathroom

B. Meal Preparation

\*1. Use Range - able to set controls; place pot or pan on at least one burner; aware of safety precautions

\*2. Use Oven - able to set controls; open and close oven door; pull out and push in racks; insert and remove pan or dish; aware of safety precautions

\*3. Use Microwave - able to set controls; open and close oven door; insert and remove dish or plate; aware of safety precautions (e.g., do not use metal containers in oven)

4. Refrigerator - able to open and close door; insert and remove items

5. Freezer - (same as refrigerator)

\*6. Dishwasher - able to open and close door; load and unload dishes; set control dial; add detergent; start machine

\*7. Garbage Disposal - able to turn on and off; has knowledge of appropriate usage; aware of safety precautions

8. Use Sink - able to reach faucets; turn water on and off

9. Wash Dishes - able to hold dishes; use cloth, brush, etc.; apply sufficient pressure to clean dishes; rinse dishes; stack to dry

10. Meal Planning

\*a. balanced meal - able to name foods that will combine to make a balanced dinner meal

- \*b. shopping list - able to plan a grocery list to buy ingredients for a balanced meal
  - c. grocery shopping
    - i. transportation - able to get to store; get home with groceries
    - ii. shopping - able to enter store; move through store; gather items; pay for items; exit store
  - \*d. follow recipe - able to give steps for following standard recipe, including listing required equipment and accurately sequencing the steps (use one of the example recipes)
11. Transport Items - able to move items such as full pots, baking trays, plates from one part of the kitchen to another (e.g., counter to sink, etc.)
  12. Pour - able to perform action of pouring liquid from one container to another (e.g., milk from carton into glass)
  13. Mix - able to hold spoon; stabilize bowl; move it through mixture in bowl (or other container) in order to combine ingredients
  14. Measure - able to accurately measure out dry and liquid ingredients in cups or spoons
  15. Peel - able to hold peeler or knife; remove peel from fruit or vegetable; aware of safety precautions
  16. Chop - able to hold knife; stabilize food on cutting surface; cut food into pieces; aware of safety precautions
  17. Open Containers - able to manipulate container and utensil (when needed) for opening; get container open
    - a. cans
    - b. cartons/boxes
    - c. jars
    - d. bottles
    - e. plastic bags
  18. Grate - able to stabilize grater; move food along grater with sufficient pressure to grate it; aware of safety precautions

## IX. COMMUNITY INVOLVEMENT

### A. Transportation

1. Bus - able to get to bus stop; board bus; pay fare; get off bus at appropriate stop; knows how to transfer
2. Paratransit - able to call to arrange for travel; give pertinent information accurately
3. Bay Area Rapid Transit (BART) - able to get to station; pay fare; get ticket; board train; get off at appropriate stop; knows how to change trains
- \*4. Map Use - able to trace route between two designated points on a street map (use example map)

### B. Financial Management

1. Savings Account - has an account; able to deposit and withdraw money
2. Checking Account - has account; able to write check; deposit and withdraw money
- \*3. Making Change - able to correctly answer the following question: "If I bought something from you and it cost \$3.25 and I gave you \$5.00, how much money would you give me in change?" (A: \$1.75)
- \*4. Budgeting - able to list at least three typical monthly expenses from the following list: rent, food, utilities (gas, electricity, telephone, water, garbage), attendant care, transportation, clothes, recreation, and entertainment

### \*C. Utilities

1. Thermostat - knows location; able to manipulate control in order to adjust temperature
2. Fuses/Circuit Breakers - knows location; knows how to determine which fuse has blown or which circuit is off; able to change fuse (using correct voltage) or move switch on circuit breaker to "on" position
3. Gas/Electricity - knows how to arrange with company to turn on and off; arrange for maintenance
4. Telephone - able to install and disconnect telephone from wall jack; able to arrange for service and maintenance

PAGE 1

CLIENT: \_\_\_\_\_  
 THERAPIST: \_\_\_\_\_  
 DATE: \_\_\_\_\_

INDEPENDENT  
 NEEDS SUPERVISION  
 NEEDS PERSONAL ASSIST (A or U)  
 ABLE TO BUT DOESN'T  
 DON'T KNOW  
 NOT APPLICABLE

WANTS TO CHANGE  
 USES EQUIPMENT

COMMENTS

I. FUNCTIONAL MOBILITY					
A. TRANSFERS AND POSITIONING					
1. Bed					
a. into					
b. out of					
c. positioning					
d. rolling					
e. lie to sit					
2. Toilet/Commode					
a. on					
b. off					
3. Bath/Shower					
a. into					
b. out of					
4. Chair					
a. on					
b. off					
c. positioning					
d. floor to chair					
e. chair to floor					
5. Car/Van					
a. into					
b. out of					
B. AMBULATION					
1. Inside					
2. Stairs					
3. Thresholds					
4. Inclines					
5. Escalators					
6. Elevators					
7. Outside					
8. Uneven Surfaces					
9. Curbs					
C. MANUAL WHEELCHAIR					
1. Self Propel					
a. inside					
b. outside					
2. Corners					
3. Thresholds					
4. Inclines					
5. Elevators					





PAGE 3

CLIENT: \_\_\_\_\_  
 THERAPIST: \_\_\_\_\_  
 DATE: \_\_\_\_\_

INDEPENDENT  
 NEEDS SUPERVISION  
 NEEDS PERSONAL ASSIST (A or U)  
 ABLE TO BUT DOESN'T  
 DON'T KNOW  
 NOT APPLICABLE

WANTS TO CHANGE  
 USES EQUIPMENT

COMMENTS

III. DRESSING									
A. UNDERGARMENTS									
1. Upper									
a. on									
b. off									
2. Lower									
a. on									
b. off									
B. UPPER BODY									
1. On									
2. Off									
C. LOWER BODY									
1. On									
2. Off									
D. FASTENING DEVICES									
1. Buttons									
2. Zippers									
3. Snaps									
4. Hooks/Eyes									
5. Buckles									
6. Velcro									
E. SHOES									
F. SHOE FASTENINGS									
G. SOCKS/STOCKINGS									
IV. FEEDING/EATING									
A. FINGERS									
B. UTENSILS									
1. Fork									
2. Spoon									
3. Knife									
C. CUP/GLASS									
V. HEALTH									
*A. KNOWS DOCTOR									
*B. KNOWS DENTIST									
C. MEDICATION									
1. Physical Ability									
*2. Manage Schedule									
*D. SKIN CARE									
*E. R.O.M.									







Appendix B

Letter to Recruit Subjects

Center for Independence of the Disabled, Inc.  
875 O'Neill Avenue  
Belmont, California 94002  
Current Date

Dear \_\_\_\_\_:

Center for Independence of the Disabled (CID) is conducting research on the Independent Living Skills Assessment Instrument that is currently being used for the First Step/Independent Living Skills Program. The results of the study will help us to improve the quality of services for CID's clients.

At this time we are contacting people who are affiliated with CID and who may be willing to be volunteer participants for this project. Volunteers should meet the following criteria:

1. Be physically disabled and over the age of 16
2. Not currently receiving First Step or Independent Living Skills services
3. Willing to give 4-6 hours of time in your home as a volunteer

The privacy of each participant will be protected throughout the study and names will not be used in the report of the results. Your help with this research would greatly benefit CID and all the clients who use our services. If you are interested in helping and/or have any questions, please call me here at CID (595-0783) by 31 October 1985.

Susy Stack-Dunne, OTR  
Coordinator  
Independent Living Skills Program

Appendix C

Consent Form



## CONSENT FORM

This is to certify that \_\_\_\_\_ has given his/her consent to participate in this study according to the following conditions:

1. That the name of the subject will not be used to identify the results of the investigation.
2. That the results will be used only for research and education.
3. That the subject has been informed of the details of his/her participation in the study.
4. That the subject is free to discontinue participation at any time during the study.

\_\_\_\_\_  
Signature of Subject

\_\_\_\_\_  
Date

\_\_\_\_\_  
Signature of Subject

\_\_\_\_\_  
Date

\_\_\_\_\_  
Signature of Subject

\_\_\_\_\_  
Date