uOttawa

L'Université canadienne
Canada's university

FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTOCTORALES

uOttawa

L'Université canadienne
Canada's university

FACULTY OF GRADUATE AND POSDOCTORAL STUDIES

Sarah Milne
AUTEUR DE LA THÈSE / AUTHOR OF THESIS

M.Sc. (Epidemiology)
GRADE / DEGREE

Department of Epidemiology and Community Medicine
FACULTÉ, ÉCOLE, DÉPARTEMENT / FACULTY, SCHOOL, DEPARTMENT

Development and Preliminary Validation of the Lower Extremity Activity Limitations Scale

TITRE DE LA THÈSE / TITLE OF THESIS

I. McDowell
DIRECTEUR (DIRECTRICE) DE LA THÈSE / THESIS SUPERVISOR

CO-DIRECTEUR (CO-DIRECTRICE) DE LA THÈSE / THESIS CO-SUPERVISOR

EXAMINATEURS (EXAMINATRICES) DE LA THÈSE / THESIS EXAMINERS

L. Brosseau

Y. Chen

Gary W. Slater
LE DOYEN DE LA FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTDOCTORALES /
DEAN OF THE FACULTY OF GRADUATE AND POSTDOCORAL STUDIES

Development and Preliminary Validation of the Lower Extremity activity Limitations Scale

By

Sarah Milne

Thesis submitted to the Faculty of Graduate and Postdoctoral Studies in Partial fulfilment of the requirements for the M. Sc. Degree in Epidemiology

Epidemiology and Community Medicine
Faculty of Medicine
University of Ottawa

Your file Votre référence
ISBN: 0-494-11355-3
Our file Notre référence
ISBN: 0-494-11355-3

I would like to thank my thesis supervisor, Dr. Ian McDowell for his patience and unwavering support. From the beginning, you have guided me and challenged me to think outside of the box. You have also taught me a great deal about the art of writing.

I would also like to thank Marg Bisch, Brenda Elliott and Dominique Surprenant for their dedication and support. You have never hesitated to fit me into your busy schedules. Your time and energy have been instrumental in the completion of this thesis.

Thanks to Fay Draper, Sylvie Desrochers and Rama Nair who have been there to provide encouragement and kind words.

I would also like to extend my gratitude to Gail Bowes for her time and office space. Both were greatly appreciated.

Thanks to Elizabeth Krisjansion for her assistance with the item response analyses.

Thanks to Lloyd Cowin and all of the Rehab staff at CHEO. You were a great help.

I will be forever grateful to the patients who made this work possible.

Most of all, thank you to my family and friends who have provided love and support.

List of Tables

List of Figures

ABSTRACT

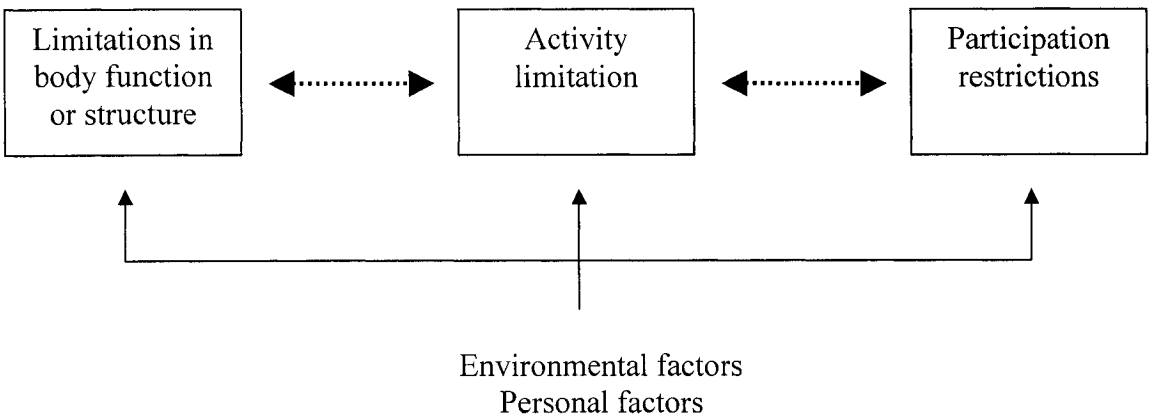
The prevalence of disabling musculoskeletal conditions in children underscores the importance of measuring function. The objectives for the current study were: 1) to develop an evaluative measure assessing function in adolescents with lower extremity orthopedic conditions and; 2) to provide a preliminary assessment of the validity, reliability and measurement properties of the newly developed scale. The Lower Extremity Activity Limitations Scale (LEALS) incorporates a 17-item self-report questionnaire and a 10-item functional performance test. Excessive internal consistency was established for both subscales of the LEALS. The construct validity of the LEALS was supported by its convergent validity with other measures of the same construct. Lastly, the LEALS self report scale was found to be more sensitive to change then the Activities Scale for Kids. Preliminary assessments suggest that the LEALS is a valid, reliable and responsive outcome measure assessing lower-extremity function in adolescents.

BACKGROUND

At a time when hospitals continue to streamline their services and there are increasing cuts to the health care system, there is a growing emphasis on evaluating all aspects of therapy, so there is an urgent need for evidence-based practice in rehabilitation. In 2001, Statistics Canada estimated that 155,000 Canadian children between five and 14 years old, or 4% of all children in this age group, had some form of activity limitation[1]. Musculoskeletal conditions in children can result from traumatic injury, congenital or neuromuscular disorders or changes related to growth and development[2]. Current practice in North America is to treat musculoskeletal conditions with early intervention programs. These are aimed at improving physical function or slowing deterioration in children with limitations. Physical function here refers to a child's ability to perform fundamental and complex activities of daily living (ADLs)[3]. Limitations in physical function can restrict a child's independence and ability to participate in the community. Therefore, the ultimate goal of rehabilitation is often to help children regain independence in their ADLs. However, a child's ability to perform activities of daily living in everyday situations (e.g., at home or at school etc.) is not often evaluated by therapists. This may in part be attributed to a lack of suitable functional outcome measures[4]. Functional outcome measures assess a patient's ability to perform ADLs, such as walking or climbing stairs.

The International Classification of Functioning and Disability (ICIDH-2), established in 2001 by the World Health Organization (WHO), is a conceptual framework which provides a multi-dimensional perspective on disability[5]. It provides a model to explore the complex process of disability, where impairments or limitations in body function or

structure interact with environmental factors to affect performance of activities of daily living and participation in society.



In recent years, dramatic improvements have been made in the treatment of children with musculoskeletal conditions[6]. However the development of outcome measurements has been unable to keep pace with advances in treatment. Pediatric research has followed the pattern of adult research, in that early studies have focused on measuring changes at the impairment level[4]. For example, studies that evaluate the change in knee range of motion following total knee replacement report on impairment outcomes (e.g. range of motion [ROM], strength). These studies are a vital starting point in the evaluation of therapeutic effectiveness but do not provide the complete picture. As the goal of many therapies is to improve activity performance in the community and reduce participation restrictions, it is essential to measure outcomes at these levels of the ICIDH-2 model. Research has also supported this extension of outcome measurement by identifying the limitations of traditional impairment measures which cannot predict limitations in activities or restrictions in participation. Several studies have demonstrated that measures of impairment only partially predict eventual function[7, 8, 9]. Therefore, if we wish to

evaluate the effectiveness of a treatment aimed at improving walking ability (an activity), it is not sufficient to measure only range of motion (an impairment) [10, 11].

In order to assess initial levels of function in ADLs and to measure change over time, reliable and valid instruments are imperative. Indeed, many methods of health measurement have merit. Self-report measures provide information on the patient's ability to do specific activities of daily living such as walking or climbing stairs and can be completed by the patient or through a proxy such as a parent or teacher or capacity to perform specific activities can be measured through direct observation. Although functional performance tests and self-report questionnaires may purport to measure the same construct, activity limitation, only weak to moderate correlations have been identified between them [3, 7, 12, 13].

Both self-report questionnaires and functional performance tests provide partial information on a patient's level of physical function. Self-reported ability to perform ADLs may reflect numerous factors including patients' performance expectations, attitudes, perceived risk or fear of pain or reinjury and psychologic distress levels [3, 14, 15] while functional performance testing is an objective measure of capacity in a controlled environment but not in real life. Although patient self-reports of activity limitation and clinician measured performance tests assess the same dimension of the ICIDH-2 model, activity limitation, both can be influenced by psychosocial and environmental factors to differing degrees. Therefore, both of these measurement approaches provide a complementary perspective on the global clinical picture. By exploring both of these perspectives and adding information from the other two levels of the ICIDH-2 framework, the clinician may be able to identify if the limitation in function is being

influenced by physical, environmental or psychosocial factors[16, 17]. But how many measures can clinicians be expected to use on a single patient?

Numerous barriers limit the widespread adoption of outcome measures into clinical practice. Common complaints include: difficulty administering the scales or scoring, lack of clinical relevance for the population of interest, time constraints and number of scales. Clinicians in orthopedic practice who treat a varied caseload are bombarded with numerous condition-specific measures. For example, in a simple search for orthopedic knee outcome measures, 13 instruments were identified [8, 18-29]. Measures were identified for general conditions of the knee [21, 23, 24]; and for specific conditions, such as patellofemoral joint disorders[18, 25, 26]; ligament deficiency [8, 19, 20, 22, 28] and joint replacement[27, 29]. Although each of these scales may be sensitive to change in a specific group of patients, the use of condition-specific scales in a varied practice would require a different scale for each diagnostic group, which presents numerous logistical and operational problems.

One approach to circumvent the need for multiple measures of health status in clinical practice is to use a more generic measure which would be applicable in a variety of conditions. However, use of a generic scale is not without its problems. A generic scale which is developed to permit comparisons across many conditions may lack specificity and precision. Although the classification of generic or specific suggests only two options, measurement instruments range across a spectrum and tools at each level of the spectrum present a different set of strengths and limitations. For the current study, at the extremes neither generic nor specific measurement tools are feasible; rather the

instrument needs be sensitive to change and applicable for a variety of lower extremity orthopedic conditions.

In summary, the prevalence of disabling musculoskeletal conditions in children underscores the importance of measuring activity limitation; the benefit of collecting information from a variety of perspectives has been established and in order to be clinically relevant, use of the outcome measure must be feasible. A lower-extremity measure that would detect change in physical function and combine information from two sources, self-report and observation would be valuable for clinicians who treat children with a variety of lower-extremity orthopedic conditions.

Literature Review

An initial literature search, of the Medline, CINAHL and Embase databases, was conducted to identify any scales, suitable for use in children, which purported to measure activity limitations in the lower extremities. From 1966 to 2003, a keyword search of "Function" or "Activities of Daily Living", or "Performance" and "Scale", or "Outcome Measure", and "Lower Extremity", or "Hip", or "Knee", or "Foot" or "Ankle" was used. In total, 68 articles were identified and reviewed for relevance. Of these, 25 articles reported on pediatric outcome measures while the remainder focused on an adult population. Those deemed relevant were critically appraised and the bibliographies reviewed for additional relevant articles. In addition to the literature review, several compendia of measuring scales were consulted [30-32].

This search of the literature identified numerous scales to measure functional limitations in adults. Fewer scales were identified for use in children and no outcome

measures on activity limitations were identified which combined self-report and observed measures of function. Considering the weak to moderate relationship between traditional impairment measures and eventual function, and self-report measures and observed measures of function, no single outcome measure will provide the clinician with a complete clinical picture. Increasingly, traditional impairment measures, self-report and observed measures of function are seen as complementary rather than competing measures, allowing the clinician to evaluate a broader clinical picture. Therefore, in the following section, all relevant self-report and functional performance tests will be reviewed.

Self Report Questionnaires

Numerous self-report questionnaires were identified to measure patients' perceptions of their ability to perform activities of daily living. Despite the proliferation of outcome measures of physical function in adults, few measures were identified, in the literature search, for the assessment of children [33, 34]. Because of the many differences in lifestyles, adult measures may be inappropriate or insensitive for use with children. A critical review of the literature on pediatric measures of function suggested that many scales previously used to assess children were derived from and tested in adult populations with little or no modification[4]. While eight generic outcome measures [11, 35-41] relevant to children with musculoskeletal conditions were identified, no general-purpose measure of physical function for children with lower-extremity orthopedic problems or conditions was found.

Two scales that assess global physical function in children and have been widely used throughout the literature are the Activities Scale for Kids (ASK)[11] and the Pediatric Outcomes Data Collection Instrument (PODCI)[35]. The ASK is a 30-question self-report measure designed to measure pediatric physical disability[11, 42]. It is a measure of global functional status and therefore comprises questions on both upper and lower extremity function. Two versions were developed, each covering the same activities but measuring slightly different concepts. The ASK-capability measures the activities a child "can do", while the ASK-performance records the activities that a child normally performs. The ASK can be completed by the child or by a parent and has demonstrated almost perfect agreement between these two ratings[43]. Although the ASK has been shown in previous research to be valid and reliable, it has been designed to cover a wide range of musculoskeletal disorders and therefore, any questions regarding upper extremity function are irrelevant for lower-extremity conditions. In addition, while it provides information on the child's perception of their ability to perform ADL it does not measure actual performance.

The Pediatric Outcomes Data Collection Instrument is a self-report questionnaire which collects information on co-morbidity, upper extremity and physical function, transfers and basic mobility, sports and physical function, pain, treatment expectations, happiness, satisfaction with symptoms and global function[35, 44]. The PODCI is available in three formats: a parent-report format for children (ages birth to ten years), a parent-report format for adolescents (ages 11 to 18 years) and an adolescent-report format for adolescents. It contains 108 items and takes an average of 10-12 minutes for the adolescent to complete and 15- 18 minutes for the parent[42]. The PODCI is a measure of

perceived functional ability that does not include objective measures of actual performance.

The literature search also identified an outcome measure developed for adults with lower extremity musculoskeletal dysfunction. The Lower Extremity Functional Scale (LEFS) is a self-report questionnaire consisting of 20 items related to activities of daily living[45]. Patients are asked to rate their ability to perform specific activities considering their lower limb problem. However, as the measure was designed for adults, many of the questions would not be relevant for the current population of interest.

Although all three scales described above measure one aspect of physical function, no one scale is entirely appropriate for the population of interest and purpose of measurement. Both the ASK and the PODCI were intended to cover a wide range of musculoskeletal disorders and therefore contain numerous questions that may be irrelevant for adolescents with lower extremity orthopedic conditions. In addition, as they try to cover numerous areas of difficulty such as upper extremity function and lower extremity function there are fewer questions which allow discrimination among different levels of lower extremity disability. The LEFS is a specific measure for lower extremity function but it was designed for adults and therefore many of the questions are not relevant or discriminating in a pediatric population.

Observed Measures of Function

In an attempt to provide an objective measure of lower extremity function, functional performance tests designed to simulate the stress placed on the body during activity in a controlled environment, have been developed. Functional performance tests are

increasingly popular because they require minimal space, equipment, time and personnel for their administration in the standard clinical context [8]. However, despite their increasing popularity, no functional performance test was identified which was developed or tested in a pediatric population. In the following section, the relevant FPT identified in the literature search will be reviewed.

The Lower Extremity Functional Test is a timed sequence of eight skills performed continuously[46]. The participant moves through each station in a preset order performing one of the eight skills: forward and backward running, side shuffling, carioca, figure-8 running, 45 and 90 degree cutting, and 90 degree crossover cutting. The LEFT was designed to incorporate many of the stresses placed on the legs during a sport activity and allow performance examination under conditions of fatigue. It is intended to help clinicians determine when it is safe for an athlete to return to competitive activity. As such, initial studies on the scales measurement properties were performed on university athletes. All eight tasks are advanced skills which require significant strength, balance and range of motion. Furthermore, when the tasks are combined in sequence they require considerable muscle endurance. Although some individual items in the LEFT might be useful, the scale surely would show floor effects and have difficulty detecting differences in functional status among adolescents more severely affected by their condition.

In addition to the generic lower extremity functional test reviewed above, many condition specific functional tests were identified [16, 17, 47-52]. Nine studies examined various functional tests including hop tests, leap and jump tests, and linear sprint, agility and stair climbing tests in patients post anterior cruciate ligament repair[8]. Other studies have focused on assessing the reliability or validity of individual performance items[53-58].

Although many of these tests demonstrated excellent reliability and validity and measure an aspect of lower extremity function, no single test covers a range of lower extremity ADLs and is appropriate for the population of interest. The scales were either developed with a specific condition or injury in mind and therefore may not be sensitive to a wider range of conditions, or are single items which may useful in part but by themselves cannot detect change over time in a wide range of conditions and severity levels.

At this point, the usefulness of a lower-extremity specific outcome measure assessing the "activity limitations" dimension of the ICIDH-2 framework has been established. A thorough review of the literature was unable to identify an appropriate tool which would combine information on environmental and personal confounding factors. Therefore, the development of a new functional scale was undertaken.

Classification of Measurement Instruments

Prior to proceeding with the development of a new scale, it was necessary to clarify the purpose of the scale. Health measurement instruments can be classified in several different ways, of which perhaps the simplest classification is based on the purpose of the instrument or type of research question that it answers: this is considered a functional classification[30]. According to Bombardier and Tugwell, scales can be classified as diagnostic, prognostic or evaluative[59]. Diagnostic measures are used primarily as assessment tools and are judged for their correspondence to a clinical diagnosis. Prognostic measures classify people according to preset criteria and can be used in screening or to predict the likelihood of a future event. Finally, evaluative measures serve as assessment tools which are designed to detect change over time.

Health measurements can also be classified descriptively according to their scope. Generic instruments are commonly used in descriptive epidemiological studies and permit comparisons across disease categories while specific instruments are generally designed for use in a clinical setting and are intended to be sensitive to changes in a person's condition over time.

Lastly, health measurements can be classified methodologically. Instruments can be developed as rating scales or questionnaires. The information can be gathered subjectively or objectively, and results can be presented as a single score (a health index) or as a profile.

For the current study, the instrument is intended to be used in a pediatric outpatient orthopedic practice to detect changes in lower-extremity function over time. For this reason, the scale will be used primarily as an evaluative tool to monitor absolute change in physical function. The scale will not be condition specific; rather it will need to be sensitive to a variety of lower-extremity orthopedic conditions. Methodologically, it will be composed of two parts: a patient self-report questionnaire providing information on the participant's subjective assessment of their function; and a functional performance test recording objective data on their capacity. Results will be presented as a profile: one score for the self-report questionnaire and a separate score for the functional performance test.

OBJECTIVES

The overall purpose of this study is to develop a measurement instrument intended to measure change in function in an adolescent population with lower extremity orthopedic conditions. The study was carried out in two phases. The first phase focused on the development and pilot testing of the Lower Extremity Activity Limitations Scale (LEALS) while the second phase provided a preliminary assessment of the validity, reliability and measurement properties of the newly developed scale.

Overall Objective:

1. To develop an evaluative measure assessing function in adolescents aged 10- to 18-years with lower extremity orthopedic conditions.

Objectives for Phase I: Development of the Lower Extremity Activity Limitations Scale

1. To construct the Lower Extremity Activity Limitations Self-Report (LEALS- SR) and Functional Performance (LEALS- FP) Scale.
2. Pilot testing
3. Revise Scale

Objectives for Phase II: Evaluation of Measurement Properties

1. Evaluate measurement properties of the scale.
2. Refine scale.

PHASE I – DEVELOPMENT OF THE LOWER EXTREMITY ACTIVITY LIMITATIONS SCALE

Item selection is an iterative process which involves many steps. Often, the results of one step influence subsequent methodologic decisions. For that reason the methods and results will be combined in the following section.

Item Selection

Reviewing existing questionnaires and surveying clinicians generated an initial pool of items to be considered for inclusion in the scale. Items were divided into two groups: questions to be included in the self-report questionnaire and performance items to be included in the observed measure of function. For the purpose of this study, content validity was established using a conceptual framework that was developed by clinical experts (S.M. and three pediatric orthopedic physiotherapists). The framework was used as a guide for item selection and outlined the areas of assessment and ADLs considered important for a lower-extremity functional assessment (Appendix 1). For example, walking ability was deemed an important area of assessment; therefore it was included in the framework.

Questions from the LEFS, ASK and POSCI formed the initial pool of items to be considered for inclusion in the self-report questionnaire. Additional questions were produced by surveying clinicians to reflect activities commonly performed by children such as running, jumping and skipping. In total, 48 items were included in the initial pool of self-report items (Appendix 2).

An initial pool of items to be considered for inclusion in the functional performance test was identified by reviewing the existing literature from sports medicine and

orthopedic rehabilitation. Twenty-four items formed the initial pool of functional performance test items (Appendix 3).

The pools of items were then reviewed for relevance by a panel of experts consisting of the primary investigator (S.M), three pediatric orthopedic physiotherapists and a survey methodologist. Initially items were reviewed for their clarity, face validity and clinical relevance. When two or more items pertaining to the same activity were identified, the most appropriate item, in terms of language, clinical relevance etc., was retained by the panel of experts. Minor modifications to the wording of individual items were made when necessary to improve the application to children. Most scales which are developed for the general public aim for a reading level of a 12 year old; however, as the current scale was intended for use in children as young as ten years old some modifications were required. For example, the LEFS refers to the participants' ability to walk a mile. Children may not be familiar with a mile, so the wording was modified to increase the relevance for the target population. Ambiguous or confusing items were eliminated. Following item review for clarity, the face validity of each item, which examines whether or not the instrument appears to be assessing the proposed construct, was examined[60]. Items that are not related to the content introduce error in the measurement, in that they discriminate among patients by a factor other than the one allegedly measured by the test. Among potential users, the scale may have difficulty gaining acceptance if it does not seem appropriate while respondents may object to an item or omit it entirely if it appears to be irrelevant. Lastly, the clinical relevance and ease of application (time, equipment and space requirements) of each item was reviewed. Each item was examined using the above criteria and the final selection of items was based on

consensus from the panel. Those items not accepted by all members of the panel were subsequently eliminated.

Following the screening of self-report items for interpretability, face validity and clinical relevance, 26 items remained, shown below in Table 1.

Table 1: 26 items included in the initial version of the LEALS – SR

Today do you or would you have any difficulty with.....
1. Putting on my socks or shoes
2. Squatting
3. Any of my usual school activities not including gym
4. My usual recreational or sporting activities including gym class
5. Lifting an object, like my backpack off the floor
6. Doing my usual jobs or chores at home
7. Getting into or out of a car
8. Getting into or out of a bus
9. Getting through heavy doors
10. Walking around inside the house
11. Walking 2 blocks
12. Walking for 30 minutes
13. Going up or down a flight of stairs
14. Climbing several flights of stairs
15. Stretching to reach a high shelf
16. Standing for 1 hour
17. Sitting for 1 hour
18. Sitting on the floor
19. Running on even ground
20. Running on uneven ground
21. Making sharp turns while running fast
22. Walking on rough or slippery surfaces
23. Walking up a gentle hill or slope
24. Walking in a crowded area
25. Hopping
26. Keeping up with my friends at school or at home

The same review of the functional performance test items resulted in the removal of 13 items. In addition, the single item pertaining to balance was divided to form two items: balance with the eyes open and proprioception assessed with the eyes closed. Gait

speed was also divided into two items: a timed 22.5-meter walk and a timed 22.5-meter run.

Following the initial review, the items remaining in the two pools were examined for content validity. The concept of content validity is closely related to that of face validity in that it assesses the coverage of the items to determine whether the scale appears to be appropriate for the intended purpose. Examination of content validity ensures that the scale has enough items to adequately cover the topics of interest. The review for content validity ensured that at least one item evaluated each of the activities deemed important in the conceptual framework. The panel of experts were supplied with a printed copy of the remaining items arranged into categories based on the conceptual framework. The items were then arranged in order of difficulty by the panel through a consensus process. When the ordering of items was complete, each panel member was asked to independently rank the items in terms of the perceived value of the information they would provide. A range of items which evaluated the ADLs deemed relevant, distributed along the continuum of disability were desired. The final selection of items was obtained through consensus.

The process of assessing content validity by the expert panel resulted in the removal of five items from the self-report scale: 9, 15, 18, 23, and 24. This resulted in a list of 21 items which were subsequently organized into similar groups of skills and re-numbered. The final scale contained 21 items, listed in Table 2.

Table 2: 21 items remaining in the LEALS – SR following a review for content validity

Today do you or would you have any difficulty with.....
1. Walking around inside the house
2. Putting on my socks or shoes
3. Squatting
4. Lifting an object, like my backpack off the floor
5. Doing my usual jobs or chores at home
6. Getting into or out of a car
7. Getting into or out of a bus
8. Walking 2 blocks
9. Walking for 30 minutes
10. Going up or down a flight of stairs
11. Climbing several flights of stairs
12. Standing for 1 hour
13. Sitting for 1 hour
14. Running on even ground
15. Running on uneven ground
16. Making sharp turns while running fast
17. Walking on rough or slippery surfaces
18. Hopping
19. Keeping up with friends at school or at home
20. Any of my usual school activities not including gym
21. My usual recreational or sporting activities including gym class

The process of assessing for content validity resulted in the removal of three items from the functional performance test. The final scale is displayed in Table 3, and contained ten items arranged by ease of administration.

Table 3: 10 items included in the final version of the LEALS – FP

1. Deep knee bend
2. Repeated sit to stand
3. Vertical jump
4. Standing hop
5. Step up
6. 12-meter hop test
7. Balance
8. Proprioception
9. 22.5 meter walk
10. 22.5 meter run

Pilot Testing and Calibration

Following scale development, pilot testing was performed on both components of the newly developed scale. The purposes of the pilot testing were to identify any difficulties with administration, to assess item clarity in the self-report questionnaire and to examine the inter-rater reliability of the functional performance subscale with the target population. Prior to pilot testing all three physiotherapists attended a mandatory training session. The objectives of the training session were to familiarize the physiotherapists with the outcome measures to be used in the study and to review patient recruitment and data collection procedures. In addition to the training session, a comprehensive operations manual was produced by the primary investigator to outline the procedures for patient recruitment, administration of the outcome measures and data collection.

Standard item clarity testing was performed for the self-report questionnaire with a convenience sample of ten patients from the Outpatient Physiotherapy Clinic at the Children's Hospital of Eastern Ontario (CHEO). A separate sample, of the same size, was obtained for inter-rater reliability testing. All patients participating in the pilot testing were identified through the Outpatient Physiotherapy Clinic at CHEO. Approval for the study was obtained from the Research Ethics Board of the Children's Hospital of Eastern Ontario. Potential participants were identified by their therapist and were considered eligible for participation if they presented with a lower-extremity condition and were aged between 10 and 18 years. Children with diagnosed cognitive impairments or limited knowledge of English were excluded because they may not have been capable of completing the self-report questionnaire. In addition, children with an associated neurological or cardiovascular condition were excluded. At the time of identification, the

referring physiotherapists provided a global rating of each child's disability level, their age and diagnosis. Informed consent was obtained by the participant and legal guardian if necessary, prior to commencement of the testing. Both samples were monitored throughout to ensure that a variety of conditions and severity levels were obtained. In addition, children from each of the nine years of the age range were recruited.

Self-Report Questionnaire

Item clarity testing was first introduced by Nuckols to ensure that all items are understood, unambiguous and jargon-free[61]. The goal of item clarity testing is to ensure that the scale is understandable to a wide range of patients. The techniques used in this study involved patients responding to the questionnaire and then explaining to the interviewer the thought process that led to their response of selected items. If this thought process indicated that the respondent had correctly interpreted the item, the item was approved.

To begin the item clarity testing the patient read the introduction to the self-report questionnaire. The introduction outlined the subject of the questionnaire and instructed the patient to check with their physiotherapist if they had any difficulties filling in the questionnaire. After reading the introductory paragraph, each patient was asked to complete the questionnaire (shown below in Table 4) to the best of their ability.

Table 4: Self-Report Questionnaire used for Item Clarity Testing

When people are sick or not feeling well it is sometimes difficult for them to do their regular activities. Please answer the following questions about your activities. If you have any questions while you are filling in the form, please ask your therapist. When you have finished filling out the form please hand it in to your therapist. Thank you for your time.

Today do you or would you have any difficulty with.....	**I could not do this at all**	**With a big problem**	**With a moderate problem**	**With a little bit of a problem**	**I could do this with no problem**
Walking around inside the house	0	1	2	3	4
Putting on my socks or shoes	0	1	2	3	4
Squatting	0	1	2	3	4
Lifting an object, like my backpack off the floor	0	1	2	3	4
Doing my usual jobs or chores at home	0	1	2	3	4
Getting into or out of a car	0	1	2	3	4
Getting into or out of a bus	0	1	2	3	4
Walking 2 blocks	0	1	2	3	4
Walking for 30 minutes	0	1	2	3	4
Going up or down a flight of stairs	0	1	2	3	4
Climbing several flights of stairs	0	1	2	3	4
Standing for 1 hour	0	1	2	3	4
Sitting for 1 hour	0	1	2	3	4
Running on even ground	0	1	2	3	4
Running on uneven ground	0	1	2	3	4
Making sharp turns while running fast	0	1	2	3	4
Walking on rough or slippery surfaces	0	1	2	3	4
Hopping	0	1	2	3	4
Keeping up with friends at school or at home	0	1	2	3	4
Any of my usual school activities not including gym	0	1	2	3	4
My usual recreational or sporting activities including gym class	0	1	2	3	4

Following completion of the questionnaire, each patient was asked a series of questions in an interview by the primary investigator. The first question asked if the respondent understood the instructions at the beginning of the questionnaire and if there were any aspects of the instructions that were not clear. The second question asked if the respondent could differentiate between response categories. This was accomplished by

testing each interval of the response scale. In order to test the first interval, respondents were asked "can you see a difference between the following responses – "I could not do this at all" and "with a big problem". This process was repeated for all five of the response categories shown below.

I could not do this at all	With a big problem	With a moderate problem	With a little bit of a problem	I could do this with no problem

Lastly, respondents were asked a series of questions designed to rate their understanding of each item. Five items were randomly selected for each participant. First, respondents were asked to explain their understanding of the activity described in the item. Then the respondent was asked if there was any part of the item or response coding that was unclear. This process was repeated for all five items for each respondent. Finally, following the structured questions, respondents were given the opportunity to add any comments they wanted.

Each item was assessed by at least two participants. The process of item clarity testing for the self-report questionnaire revealed no inconsistencies or ambiguities relating to the instructions, method of scoring or the scale items themselves. All items were understood by all ten patients, and no problems with language were identified. One problem with formatting was identified during pilot testing. Several subjects were observed having difficulty identifying the response categories for individual questions. As a result, shading was introduced for every second question to facilitate identification of the corresponding response set. None of the items were removed as a result of the item clarity testing with a subset of the target population. The Lower Extremity Activity

Limitations Scale – Self Report (SR) questionnaire (Table 5) was deemed ready for validity testing.

Table 5: Lower Extremity Activity Limitations Scale, Self-Report (SR)

When people are sick or not feeling well it is sometimes difficult for them to do their regular activities. Please answer the following questions about your activities. If you have any questions while you are filling in the form, please ask your therapist. When you have finished filling out the form please hand it in to your therapist. Thank you for your time.

Today do you or would you have any difficulty with . . .	I could not do this at all	With a big problem	With a moderate problem	With a little bit of a problem	I could do this with no problem
Walking around inside the house	0	1	2	3	4
Putting on my socks or shoes	0	1	2	3	4
Squatting	0	1	2	3	4
Lifting an object, like my backpack, off the floor	0	1	2	3	4
Doing my usual jobs or chores at home	0	1	2	3	4
Getting into or out of a car	0	1	2	3	4
Getting into or out of a bus	0	1	2	3	4
Walking 2 blocks	0	1	2	3	4
Walking for 30 minutes	0	1	2	3	4
Going up or down a flight of stairs	0	1	2	3	4
Climbing several flights of stairs	0	1	2	3	4
Standing for 1 hour	0	1	2	3	4
Sitting for 1 hour	0	1	2	3	4
Running on even ground	0	1	2	3	4
Running on uneven ground	0	1	2	3	4
Making sharp turns while running fast	0	1	2	3	4
Walking on rough or slippery surfaces	0	1	2	3	4
Hopping	0	1	2	3	4
Keeping up with friends at school or at home	0	1	2	3	4
Any of my usual school activities not including gym	0	1	2	3	4
My usual recreational or sporting activities including gym class	0	1	2	3	4

Functional Performance Test

Reliability, or the degree of consistency of a measure[62, 63], is influenced by five sources of variance: inter- and intra-subject, inter- and intra-observer and error. Pilot testing of the functional performance subscale was undertaken to evaluate two sources of variance: inter-subject variability and inter-observer variability. Several different therapists would be conducting the performance testing; we therefore wished to evaluate the inter-rater reliability or agreement between the observers. If inter-rater reliability is found to be high then it is not essential to assess the intra-rater reliability [60]. Prior to inter-rater reliability testing, instructions for administration and scoring were reviewed with all of the therapists.

In order to measure the inter-rater reliability, three physiotherapists were asked to simultaneously observe a sample of children performing the items on the subscale. To begin the pilot testing, I reviewed the purpose of the testing and introduced the participant to the evaluators. Following the introductions, I lead the participant through each item of the performance scale using the standard process described in the operations manual. Therapists were asked to independently score and record each item on the standardized evaluation form shown below (Table 6).

Table 6: Lower Extremity Activity Limitations Scale, Functional Performance (FP)

Children's Hospital of Eastern Ontario Centre hospitalier pour enfants de l'est de l'Ontario	**NOTES**
1. Double leg squat performed on a flat surface 0 = unable 1 = much difficulty (less than ½ knee ROM) 2 = some difficulty (more than ½ knee ROM but not full OR with pain) 3 = able	_____ _____ _____

2. Repeated sit to stand with reach
   Subject is required to transition from sitting on a chair to standing and touching their dot, without using their hands. Count the number of repetitions performed in 30 seconds.

   Number of Repetitions _____

3. Two-foot static vertical jump

   Height of child _____ cm

   Height of jump _____ cm

   [Height of jump – Height of child] _____ cm

4. One-legged standing hop

   Distance _____ cm

5. Step up
   Affected leg remains on a 10 inch step while the other leg steps up and down. Count the number of repetitions performed in 30 seconds.

   Number of Repetitions _____

6. 12-meter hop test

   Time to complete _____ seconds

7. One-legged static balance on affected leg
   Subject required to stand on affected leg with his/her **eyes open**. Time recorded. Maximum 30 seconds.

   Time _____ seconds

8. One-legged static balance on affected leg
   Subject required to stand on affected leg with his/her **eyes closed**. Time recorded. Maximum 30 seconds.

   Time _____ seconds

9. 22.5 meter walk

   Time to complete _____ seconds

10. 22.5 meter run

    Time to complete _____ seconds

Functional performance test forms were collected immediately following completion of the test by each participant. The process was repeated until a sample of ten children had been evaluated.

Following completion of the initial inter-rater reliability testing it was determined that a fourth assessor should be added to carry out the performance testing if necessary (in case one of the original therapists was ill, etc.). In order to examine the inter-rater reliability of the new assessor, ten additional children were recruited. The new assessor was required to score the performance of at least three children with each of the other three therapists following the same protocol described above.

Two statistics were used to measure the inter-rater reliability. For the initial round of inter-rater reliability testing using scores from the three original physiotherapists, the nine continuous or discrete items were evaluated using an intra-class correlation coefficient. In addition, the maximum disagreement in units was calculated for all ten items. In order to assess the inter-rater reliability of the fourth evaluator, the maximum disagreement for each item was calculated.

Pilot testing revealed no difficulties with administration of the LEALS-FP. In addition, good inter-rater reliability was obtained for all ten items. For the nine continuous or discrete items, inter-rater reliability was found to be very high. The ICC ranged from 0.97 to 1.00 (95% CI: 0.92 to 1.00). The maximum disagreement for the timed items was found to be less than one second, which is consistent with values published in the literature for stopwatch reaction time. The maximum disagreement for the four discrete items was limited to two units. A summary of the inter-rater reliability results for the

continuous and discrete items is presented in Table 7. Detailed scoring for Item 4 (Standing hop) of the LEALS-FP, by the initial three observers is provided in Table 8.

Table 7: Intraclass correlation coefficients and maximum disagreement values for the nine continuous or discrete performance test items

		ICC (SEM)	Difference*Δ □		
			Mean	Min.	Max.
2.	Repeated sit to stand*	0.99 (2.1)	0.40	0	2
3.	Vertical jump Δ	0.99 (1.8)	0.50	0	2
4.	Standing hop Δ	0.99 (2.9)	0.70	0	2
5.	Step-ups*	0.99 (0.81)	0.10	0	1
6.	12-meter hop test □	0.97 (0.95)	0.30	0.05	0.77
7.	Static balance □	1.00 (0.1)	0.06	0.00	0.54
8.	Proprioception □	0.99 (0.75)	0.18	0.00	0.60
9.	22.5-meter walk □	0.99 (0.74)	0.25	0.00	0.66
10.	22.5-meter run □	0.98 (0.56)	0.22	0.01	0.48

* measured in units, Δ measured in cm, □ measured in seconds

Table 8: Individual participant scores for Item 4 (Standing hop) of the LEALS- FP with associated maximum disagreement values

	Rater 1	Rater 2	Rater 3	Disagreement (cm)
1	120	120	120	0
2	84	84	84	0
3	95	96	96	1
4	90	90	90	0
5	102	103	102	1
6	78	80	78	2
7	84	85	86	2
8	190	190	190	0
9	92	92	92	0
10	94	92	94	2

For the single item with an ordinal response scale, good agreement was established between all four observers. Disagreement in scoring occurred for a single participant and varied by only one unit.

No items in the functional performance test were removed following inter-rater reliability testing. The LEALS-FP was deemed ready for validity testing.

Additional Instruments

Pilot testing or calibration was also undertaken for several additional instruments administered for the study. These evaluations were conducted using several different methods but did not require the use of a patient sample.

Additional measurement instruments were included in order to describe the patient population and to provide data for hypothesis testing and construct validitation. These additional instruments included: the Activities Scale for Kids, three quality of life questions, a pain visual analogue scale, three impairment measures [joint ROM, manual muscle testing, and flexibility] and three questions regarding the severity and impact of the patients' condition. Pilot testing or calibration was undertaken for the objective measures to be completed by the physiotherapists. These included the three impairment measures and three questions concerning the patients' condition. Pilot testing was not performed for the three instruments that would be included in the self-report portion (the ASK, three quality of life questions and the pain scale) of the study as all three instruments have been used previously in research with similar patient populations [42, 64].

Calibration of the objective measures was accomplished by conducting a two hour instrument review and training session. During the training session the methods for data collection were reviewed with all therapists. This included a review of the instruments, patient positioning, therapist positioning, and scoring scales. In addition, the standardized data collection forms were reviewed and modified as required. All of the information

presented in the instrument review session was organized into an operations manual which remained on-site for the duration of the research project.

PHASE II – Evaluation of the Measurement Properties of the LEALS

Phase two of this study involved the administration of the LEALS – SR and FP to a larger patient population in order to complete a preliminary assessment of the scales' validity, reliability and measurement properties.

METHODS

Population

All patients participating in Phase II of the study were identified though the Outpatient Physiotherapy Clinic at the Children's Hospital of Eastern Ontario. Inclusion criteria were: aged between 10- and 18- years and referred for physiotherapy with any lower-extremity musculoskeletal condition. Children with diagnosed cognitive impairments or limited knowledge of English were excluded because they may not have been capable of completing the self-report questionnaire. In addition, children with an associated neurological or cardiovascular condition were excluded.

Using a master list of all new referrals to the Outpatient Physiotherapy Clinic, eligible patients were notified of the study during an initial telephone contact. Upon arrival at the Outpatient Physiotherapy Clinic, they were introduced to the primary investigator who provided an information letter (Appendix 4) and consent form (Appendix 5) outlining the purpose and requirements of the study. If the patient agreed to participate the primary investigator obtained informed consent from the child and legal guardian if required.

Approval for the study was obtained from the Research Ethics Board of the Children's Hospital of Eastern Ontario, and each patient gave informed consent prior to participation in the study.

Sample Size

The selection of a sample size was based on the balance between the relative gains in information with an increasing sample size and the time constraints of patient recruitment. The primary statistical analysis for validity testing was to be performed by examining correlations with other accepted measures. It was anticipated that moderate correlations (0.5 to 0.7) would be demonstrated between self-report or functional performance measures and accepted measures of function, impairment and quality of life. As shown in the Table 9, confidence intervals broaden greatly at lower levels of correlations; for a correlation of 0.4, a sample size of 100 would produce a 95% CI of 0.22 to 0.52 while for a correlation of 0.6, the same sample size would result in a 95% CI of 0.46 to 0.71. Although, a sample size of 100 is somewhat underpowered to obtain a narrow confidence interval around the point estimate of low to moderate correlations, it would describe a pattern of convergent correlations with physical measures. Furthermore, the gain in precision is minor for an increase in sample size of 100 to 150. Finally, based on caseload information provided by the Outpatient Department, a sample size of 100 seemed feasible to obtain in the desired recruitment period.

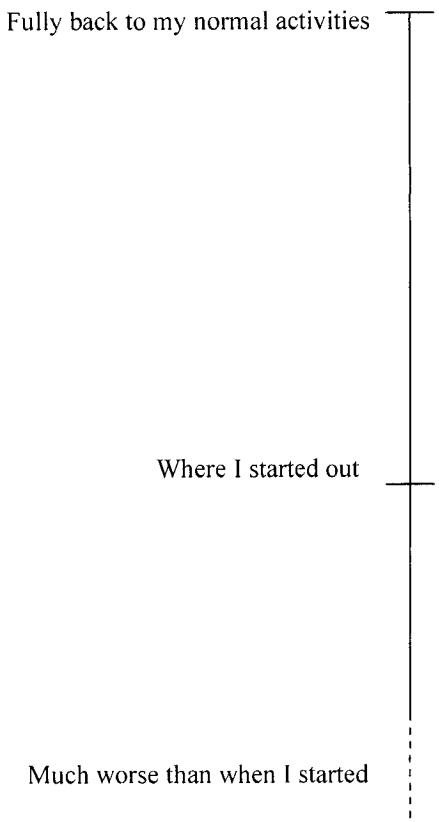
Table 9: Width of 95% confidence intervals around selected Pearson correlations, for varying sample sizes [65]

N	**0.40**	**0.50**	**0.60**	**0.70**	**0.80**
50	0.14 – 0.61	0.26 – 0.68	0.39 – 0.75	0.53 – 0.82	0.67 – 0.88
100	0.22 – 0.52	0.34 – 0.63	0.46 – 0.71	0.58 – 0.79	0.72 – 0.86
150	0.26 – 0.53	0.37 – 0.61	0.49 – 0.69	0.61 – 0.77	0.73 – 0.85
200	0.28 – 0.51	0.39 – 0.60	0.50 – 0.68	0.62 – 0.76	0.74 – 0.84

Administration of Self-Report Items

Following enrolment into the study, patients were asked to complete, on their own, a booklet consisting of two questionnaires and a selection of supplemental items (Appendix 5). In addition to the LEALS - SR questionnaire, the booklet contained a 10-point numerical rating pain scale, three questions on quality of life and the Activities Scale for Kids (ASK). This material was completed at all visits. In addition, on the initial assessment, patients were asked three questions regarding their level of physical activity. During follow-up visits, patients were asked to complete a visual analogue scale (VAS), shown below, regarding their perceived progress with treatment.

Please mark on the line below where you think you are in your recovery from this injury.



Each patient was asked to complete the study evaluation a maximum of four times. Evaluations were conducted on initial assessment and then at selected routine follow-up visits (typically three-week intervals) for six weeks or until the child was discharged (in cases where discharge occurred prior to six weeks). The measure was also administered in a sub sample of patients (at a regularly scheduled appointment) up to 72 hours following the initial administration in order to examine the test-retest reliability. As many adolescents with musculoskeletal conditions demonstrate true change over a short period of time, test-retest reliability was examined using a subgroup of patients identified as having a chronic condition.

While completing the booklet, participants were informed that the primary investigator would be nearby, if they had any questions. Demographic information (gender, age, diagnosis) was obtained directly from the patient referral record. Participants required an average of ten minutes to complete the self-report booklet.

Selection of Study Instruments

Additional measurement instruments were selected with the goal of describing the patient population and providing data for hypothesis testing in order to assess validity.

Activities Scale for Kids

The ASK was included as a widely accepted measure of physical disability in children [11 42 66]. It was incorporated into the self-report booklet in order to be used for construct validation and to allow direct comparisons of the measurement properties (ceiling and floor effects, sensitivity to change) of the two scales. We hypothesized that the LEALS-

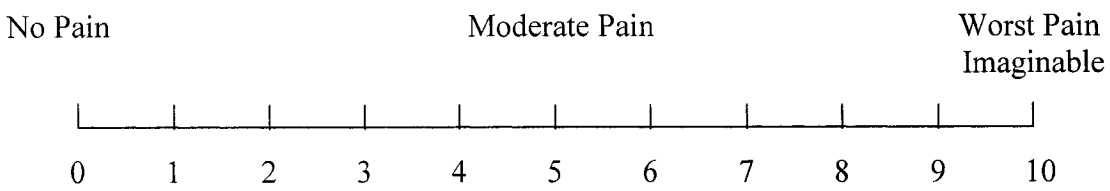
SR would be superior to the ASK at discriminating between different levels of functional ability and measuring change with treatment.

The ASK is a self-report measure that contains 30 items representing nine domains. It was designed as a measure of global functional status and has been shown in previous research to be valid and reliable[34]. Lower scores indicate more disability, with a total range of 0 to 120. Two versions of the scale are available: the ASK-capability measures what the child "could do", while the ASK-performance measures what the child usually does. For the current study, the ASK-capability was included because the investigators were interested in obtaining information on the child's perception of their ability. In addition, the newly developed scale examines the child's perception of their ability to perform specific activities of daily living.

Pain

A numerical and verbal rating (shown below) was also included in order to assess level of pain.

On average over the past two days, how much has your injury hurt?



It was hypothesized that pain would affect a patient's ability to perform functional activities. Those patients who presented with more pain would have greater difficulty completing the functional tasks and would therefore have lower scores. The decision to

use a numerical and verbal scale was made for several reasons. Firstly, numerical and verbal rating scales have been used extensively in medicine to measure pain[67]. They are also very simple and quick to administer. Lastly, the scale used for this project was developed and tested at the Children's Hospital of Eastern Ontario in a pediatric population.

Quality of Life

Three questions on quality of life were also included for the purpose of construct validation. It was hypothesized that patients with a higher perceived quality of life would score higher on the evaluation of function than similar patients with a lower perceived quality of life. The following three questions from the adolescent version of the COOP charts, a generic quality of life scale with demonstrated validity and reliability in a pediatric population[68], were selected: During the past month, how often did you feel anxious, depressed, irritated, sad or downhearted and blue?; During the past month, if you needed someone to listen or to help you, was someone there for you?; During the past month, how often did you talk to your problems, feelings or opinions with someone on your family?

Administration of Observed Measures of Function and the LEALS-FP

Following completion of the self-report booklet, participants were introduced to their physiotherapist who conducted the physical assessment. All three physiotherapists involved in the study were experienced in assessing and treating adolescents with musculoskeletal conditions and were working in the Outpatient Physiotherapy Clinic for the duration of the study.

Data collection for the study was conducted during routine appointments at the same time as participants were being evaluated or treated for their condition. Therapists were allowed to evaluate measures of impairment (ROM, strength, and flexibility) at any time during their assessment but prior to any treatment. Impairment measures were performed using a mutually agreed upon standardized approach and recorded on standard assessment forms (Appendix 7). Therapists were only required to take measurements around the affected joint on both legs. For example, if the participant presented with a knee injury, therapists were required to take range of motion, strength and flexibility measurements at both knees. However, they were not required to record information at the level of the hip or ankle. The impairment measures required approximately ten minutes to complete.

Following the assessment portion of the appointment, the ten-item functional performance test (Appendix 8) was administered by any one of the four assessors who participated in the reliability testing. Items were performed in order, using standard commands. All equipment required for the testing was set up prior to recruitment and remained consistent for the duration of the study. All timed items were measured using a

calibrated stop watch. Items which required a distance or height reading were equipped with a metric scale.

Selection of Study Instruments

Three impairment measures were used for the construct validation of the LEALS-FP. In the absence of a gold standard to measure lower-extremity function, construct validity assessed the extent to which the performance scale correlated with the following subset of impairment measures belonging to the same construct.

Range of Motion

Range of motion at each joint in the lower extremity was assessed following the standardized approach described by Clarkson and Gilewich[69]. A goniometer was used to take a precise reading at each joint. Standard positioning of the patient, therapist and measurement tool were used. This method of measurement was selected because it is widely used in rehabilitation; it has been shown to be valid and reliable in numerous studies and is relatively quick and easy to perform.

Strength

Muscle strength surrounding the affected area was assessed using manual muscle testing (MMT). The patient's maximum muscle strength through one repetition was assessed using the five-point grading scale developed by the Medical Research Council[70]. Testing was performed following the standard approach for patient and therapist positioning described by Kendall[71]. MMT was chosen because it is widely used in clinical practice; it

does not require any equipment; and has been shown to be valid and reliable when performed by trained clinical evaluators[72].

Flexibility

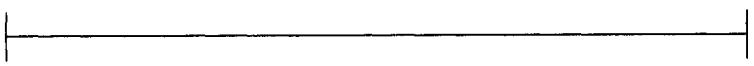
Flexibility in the surrounding muscles was measured in degrees using a goniometer. Each muscle was evaluated using the standard approach for patient and therapist positioning described by Kendall[68].

Supplemental Items

These additional items were included to input into a regression analysis in order to determine what proportion of the variance observed in the scale scores (LEALS-SR and FP) can be attributed to the various predictor variables.

Impact VAS

A 10 cm VAS was included to allow the therapists to rate the impact of the child's condition on their ADLs. Therapists were asked to determine "based on their experience with injuries of this nature, what impact would you expect this injury to have on the child's current activities of daily living?" The scale was anchored with "Severely Disabling" and "Minimal Impact" at the extreme ends.



Severely Disabling Minimal Impact

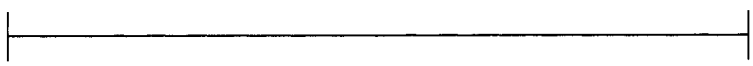
Response VAS

A 10 cm VAS was also included to allow therapists to rate the child's response to their condition. Therapists were to rate, based on their interaction with the patient, how they felt the patient was dealing with their condition.

In denial Appropriate response Exaggerated response

Improvement VAS

Lastly, a 10 cm VAS was used to measure change at follow-up visits. The improvement VAS was not included for the initial assessment. Therapists were asked to rate the degree to which the patient's condition had changed since the initial assessment.



Much worse Worse No change Improvement Large Improvement

Data Collection and Entry

As three physiotherapists were required to collect data on study subjects, there was considerable potential for data collection errors. Missing data was minimized by ensuring that all three physiotherapists used standardized forms for each assessment. In addition, data collection forms were retrieved immediately following the assessment and reviewed for completeness. To minimize the occurrence of incorrect data, all data entered into the database was double checked by the primary investigator (S.M). In addition, original forms were stored in a central location and were available to cross-check data. Data collection procedures were standardized and data quality was monitored for completeness, internal consistency and timeliness. As a further precaution, checks were built into the data entry program to protect against impossible values (e.g. ankle dorsiflexion > 40 degrees, strength >5/5)

Analyses

Response Patterns

The first step was to examine the distribution of the responses in the overall scale and for each item in the Lower-Extremity Activity Limitations Scale: self-report questionnaire and functional performance test. Items with extremely skewed response patterns (either very low or very high) were reviewed. A skewed distribution may result in ceiling or floor effects, where most of the responses are clustered at the extreme high end or low end of the scale. When ceiling effects occur, a scale cannot detect improvement, or distinguish among positive responses, whereas floor effects make it difficult to distinguish between respondents at the lower end of the underlying construct. When too

many children score near the highest or lowest possible values on a scale, the scale is unable to distinguish among the desired range of health states.

One method for examining response patterns is to examine the endorsement frequency, or proportion of people who give each response option to an item. With a 5-point Likert scale as used in the self-report questionnaire, each item has five "frequencies of endorsement": the proportion choosing option A, the proportion choosing option B etc. Streiner and Norman suggest that endorsement rates should not be greater than 0.80 for any single response category [59]. Options with high endorsement rates (>80%), where most people are responding with the same alternative, do not improve the scale's measurement properties yet make the scale longer. These items should therefore be eliminated. For the current analysis, items showing greater than 80% of respondents in a single response category were rejected.

Reliability

Internal consistency coefficients assess how well items measure an underlying construct. The current scale was developed with the objective of measuring function in the lower extremities; therefore we anticipated that each item included in the scale would measure an aspect of function. In order to test this assumption we examined the internal consistency or homogeneity of the scale. If all of the items in the scale are in fact measuring an aspect of physical function, then items should be correlated with each other. However, items should not be so highly correlated that they are redundant. Cronbach's α is a measure of reliability through correlation – it reflects the average correlation among items of a scale, taking into account the total number of items in the

scale. A low α may suggest that the scale is measuring several different concepts while a very high α may suggest item redundancy. Standards for internal consistency have been described by Nunnally and Bernstein and suggest that coefficients should be above 0.70 to reassure that the scale measures a single concept, but not higher than 0.90 which suggests redundancy and needless length[73]. If a scale is homogeneous then all items should also correlate with the total score (item-total correlations).

Construct Validity

The concurrent validity of the LEALS – SR was examined by correlating the total score with the Activities Scale for Kids. The ASK differs from the LEALS in that it was developed for children with a broad range of musculoskeletal conditions. In contrast, the LEALS is intended to be used only in children with lower-extremity orthopedic conditions.

In the absence of an accepted criterion measure, determination of the validity of new scales relies heavily on the concept of construct validity. As an initial assessment of the construct validity, this study focused on the convergent validity of the LEALS with other measures of the same construct. One or more predictions are established in advance and the extent to which a measure yields results concordant with the anticipated results provides support for the validity of the measure. In this study, we believed that validity for our measure would be supported if:

1. A moderate to high correlation ($r > 0.6$) would exist between self-report LEALS scores and ASK scores.
2. A low to moderate correlation would exist between lower-extremity function and impairment measures (range of motion, strength, flexibility and pain).
3. A moderate correlation would exist between the self-report and functional performance test scores of the Lower-Extremity Activity Limitations Scale.
4. Patients with more severe conditions, judged by the three physiotherapists, would demonstrate lower scores on the self-report and performance scales than patients with mild conditions.

Sensitivity to Change

Sensitivity to change was assessed by comparing the mean change in scores observed on the LEALS-SR to those of the ASK. The ASK was expected to have ceiling effects limiting its sensitivity to change, therefore, we expected the LEALS would perform better. Scale scores for the LEALS-SR and ASK were converted to a score between 0 and 100 to allow direct scale comparisons. Analyses were performed on a subsample of children who were deemed to have changed through patient and clinician prognostic ratings.

Analytic Procedures

Descriptive Statistics

Demographic information was analyzed by producing frequency tables, means, standard deviations and histograms using version 12.0 of the SPSS statistical package for Windows. Response histograms were produced to assess the overall performance of the

self-report questionnaire, functional performance subscale and individual scale items. The frequency distribution, endorsement rate, mean, and standard deviation were also examined for each item. Scatter plots were produced to compare the performance of the self-report questionnaire with the Activities Scale for Kids. In addition, ceiling and floor effects were compared.

Factor Analysis

Exploratory factor analysis was performed initially including both components of the Lower Extremity Activity Limitations Scale (SR and FP), using repeated measures data. The purpose of factor analysis is to display the correlations among a number of related variables. Variables that are highly correlated among each other are then grouped together into a "factor" such that variables within a factor are more highly correlated with variables in their factor than with variables in other factors. For the current study, factor analysis was used to confirm the separation of two scores (one for the self-report questionnaire and a second for the functional performance test) for the LEALS and to identify any items that do not correspond with the underlying construct.

For the purpose of this study it was hypothesized that the self-report scale would only be moderately correlated with the performance items. Therefore, we anticipated that self-report items and performance items would form two distinct factors supporting the presentation of LEALS scores as a profile.

Principal components analysis was used to determine the number of factors. Although principal component analysis is different from factor analysis, they often produce similar results[74]. Both factor analysis and principal component analysis derive linear

combinations of the original variables, but principal components analysis is mathematically simpler. Each component is constructed to be uncorrelated with the others. The analysis extracts components in order based on the variance they account for. The first component accounts for the largest amount of variance in the data, the second component accounts for the next largest amount of variance remaining in the data and so on until all of the variance has been accounted for.

Principal components analysis was used to produce two different factor solutions: an unrotated factor matrix and a rotated factor matrix. In general, rotation of the factors simplifies interpretation of factors. There are two major types of rotation: orthogonal rotations, where the new factors remain uncorrelated and oblique rotations where new factors are correlated. As a moderate correlation was anticipated between self-report and performance items, we decided to perform an oblique rotation.

Principal components analysis with oblimin rotation (delta = 0) was used to produce a factor matrix. To determine the number of factors we examined eigenvalues, scree plots and the ratio of eigenvalues. The size of the eigenvalue reflects the dispersion of the data in a multivariate space that has one axis for each component[75] while a scree plot is a graphical representation of successive eigenvalues. In an analysis with one clear factor, a ratio 5:1 or greater can be seen between the first and second eigenvalues.

Principal components analyses were repeated separately for the self-report items and the functional performance test items. Individual item loadings were examined to identify items in either scale that did not support the underlying construct. Item loadings of less than 0.5 were considered suspect.

Item Response Theory

Item response theory (IRT) analyses were used to assess the difficulty and discriminative ability of each item in the self-report questionnaire. IRT uses a mathematical model of the relationship between the observable responses on a test and the unobservable trait or ability level being measured[76]. IRT is based on two assumptions: data are unidimensional (i.e. the items only measure one trait); and the probability of answering any question in a positive direction is unrelated to the probability of answering any other item in a positive direction for people with the same amount of the trait. If these assumptions are met, an IRT model can be used to estimate characteristics of individual test items and place these estimates on a standardized ability scale, traditionally known as theta.

The Multilog program was used to produce Marginal Maximum Likelihood item parameter estimates for the latent trait "lower-extremity function". A two parameter graded-response model was used to describe both the level of difficulty and discriminative ability of individual test items. Repeated measures data were used for the analyses. The logistic form of the two-parameter model is defined as:

$$P_i(\theta) = \frac{1}{1 + e^{-ai(\theta - bi)}}$$

where $P_i(\theta)$ is the probability of an affirmative response for item i at ability θ. The parameter b_i is the difficulty index, defined as the point on the ability scale at which the probability of an affirmative response is 50 percent. The parameter a_i is the item discrimination index, an indicator of how sharply an item distinguishes well from sick respondents.

For each item, item characteristic curves (ICC) were plotted and the discriminating power (a) and item location (b) were calculated. In addition, the reliability or consistency with which the IRT model separated subjects from low to high ability was calculated. Test information and standard error curves were also produced for the LEALS-SR total score.

IRT analyses could not be performed on the LEALS-FP because the scale did not use categorical response categories.

Correlations

To measure the correlation between two continuous, normally distributed variables, Pearson's correlation coefficient was utilized. Results were calculated with a 95% confidence interval. Values may range from -1 to 1, where r = 1 indicates a perfect positive linear relationship, while r = -1 indicates an inverse linear relationship. If r = 0, there is no linear association between the two variables. For the current study, a high correlation was considered to exist when $r \geq 0.7$, a moderate correlation occurred when r fell between 0.4 and 0.7, and a low correlation was present if r was between 0.2 and 0.4. Correlation coefficients were produced initially using data from the initial assessment (n = 52). A second analysis was then performed using repeated measures data from all three time points (n = 122), as was done in a study by van Excel[77].

Linear Regression

Multiple linear regression using combined data from all assessments was performed to quantify the relationship between one or more of the predictor variables and LEALS

scale scores (SR and FP), while controlling for the other variables. The purposes of the regression were to characterize the relationship between the dependent and independent variables, to determine which of the independent variables was important for predicting function and to assess the interaction effects of the independent variables.

Significant variables ($p \leq 0.05$) were selected using stepwise selection. An examination of the variables selected into the model and the slope of the line was used to characterize the relationship between the independent and dependent variables. The adjusted R^2 and standardized residual plots were used to assess the fit of the model.

RESULTS

Sixty-four adolescents undergoing physiotherapy treatment were invited to participate in the study. Sixty-one agreed and three refused. Of the three individuals who refused participation, two reported they would not be returning for physiotherapy and the third declined due to time constraints. Following screening, seven individuals were judged ineligible to participate: three were unable to provide informed consent (under sixteen but not accompanied by a parent), two did not have adequate proficiency in English to complete the questionnaire; and two had cognitive impairments which limited their ability to respond to the questionnaire. Following enrolment, one patient was excluded due to an associated neurological condition. Demographic information obtained on the remaining 53 respondents is presented in the table below.

Table 10: Baseline Characteristics of the 53 Study Participants

Demographic Characteristics	Proportion
Gender (% female)	53%
Age	13.6 years (sd ± 2.1)
Medical Condition	
Hip	
Arthritis	2%
Knee/ thigh/ leg	
Patellofemoral pain	34%
Post-surgery	8%
Subluxing/ dislocating patella	8%
Fracture	4%
Muscle strain	4%
Osgood Schlatter	3%
Ligament sprain	2%
Foot/ ankle	
Fracture	13%
Ligament sprain	9%
Plantar fascitis	9%
Muscle strain	2%
Post-surgery	2%
Affected Leg	
Right	38%
Left	39%
Both	23%
Time since onset	
Acute (< 4 weeks)	9%
Sub-acute (4 – 12 weeks)	36%
Chronic (> 12 weeks)	55%
Severity of condition	
Mild	43%
Moderate	39%
Severe	18%
Activity Level	
High (5 - 7 days/ week)	38%
Moderate (2 - 4 days/ week)	46%
Low (≤ 1 day/ week)	16%

The participants in this study were slightly more likely to be female (53%) with an average age of 13.6 (sd ± 2.1) years. The majority of participants presented with a chronic musculoskeletal condition (55%) occurring in a single leg (77%). Knee

conditions accounted for 63% of the sample with the most common diagnosis being patellofemoral pain syndrome. Eighteen percent of patients were rated by their therapist as having a condition which severely affected their physical function. When asked about participation in physical activity, 46% of the participants reported a moderately active lifestyle, which involved participation in a physical activity between two and four days per week.

Data were collected over 8 months. In total, 122 complete assessments were performed on the 53 participants. A summary of the data collected at different time-points is presented in Table 11. No children received follow-up within 72 hours.

Table 11: Summary of assessments

	Complete	Discharged	Lost to follow-up
Initial assessment	53	0	0
72-hour follow-up	0	0	53
Three-week follow-up	42	5	6
Six-week follow-up	27	12	3
Total assessments	122		

Scale Statistics

The 21-item LEALS-SR produced total scores at baseline ranging from 23 to 83. No ceiling (score of 84) or floor (score of 0) effects occurred. Six percent of the sample was found to have scored in the top 5% of possible scores, while 0% was found to have scored in the bottom 5% of the total score range. The distribution of scores is shown below in Figure 1.

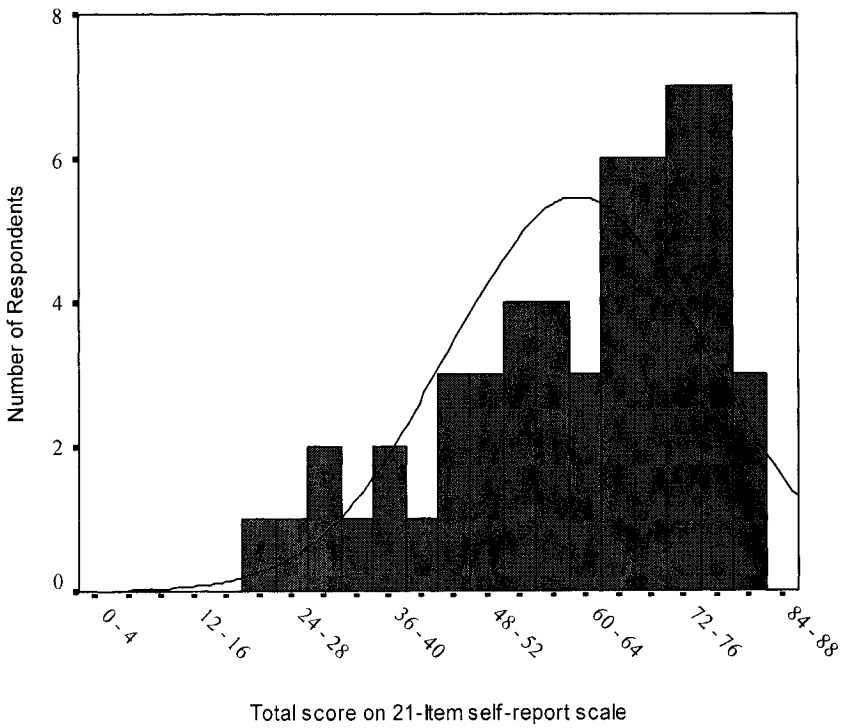


Figure 1: Distribution of Initial Assessment scores of the 21- Item LEALS-SR

The 10-item LEALS-FP produced total scores at baseline ranging from 0 to 87. The mean was 60.4 with a standard deviation of 20.2. No ceiling (score of 100) or floor (score of 0) effects occurred. The distribution of scores is shown below in Figure 2.

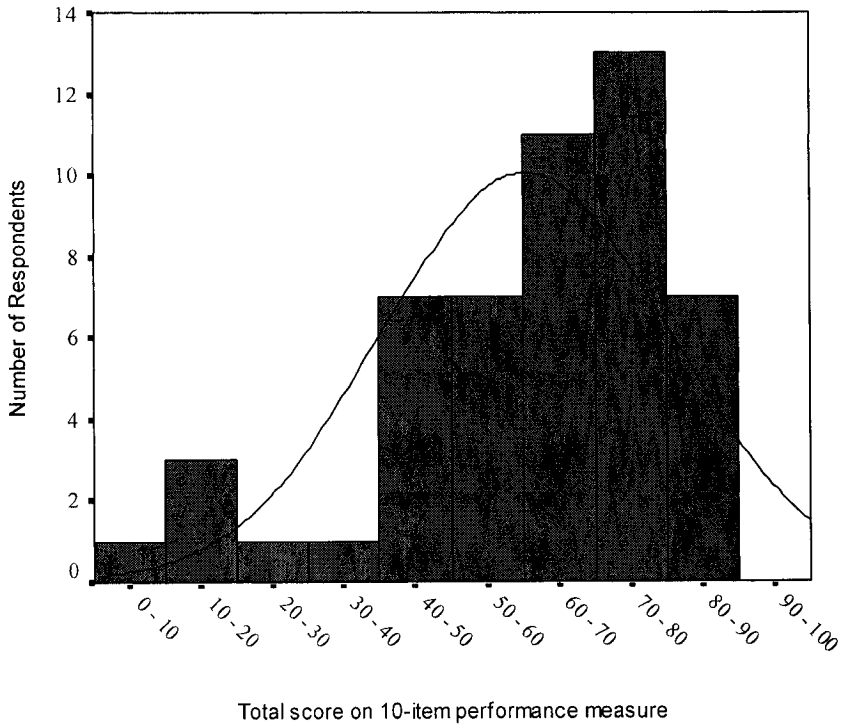


Figure 2: Distribution of Initial Assessment scores of the 10- Item LEALS - FP

The 30-item Activities Scale for Kids produced total scores at baseline ranging from 43 to 120; see Figure 3. Ceiling scores (score of 120) were found in 8% of the sample. In addition, 40% (21/53) of the sample was found to have scored within the top 5% (≥114) of the total score range. No floor effects were found and 0% of the sample scored within the bottom 5% of the scale range.

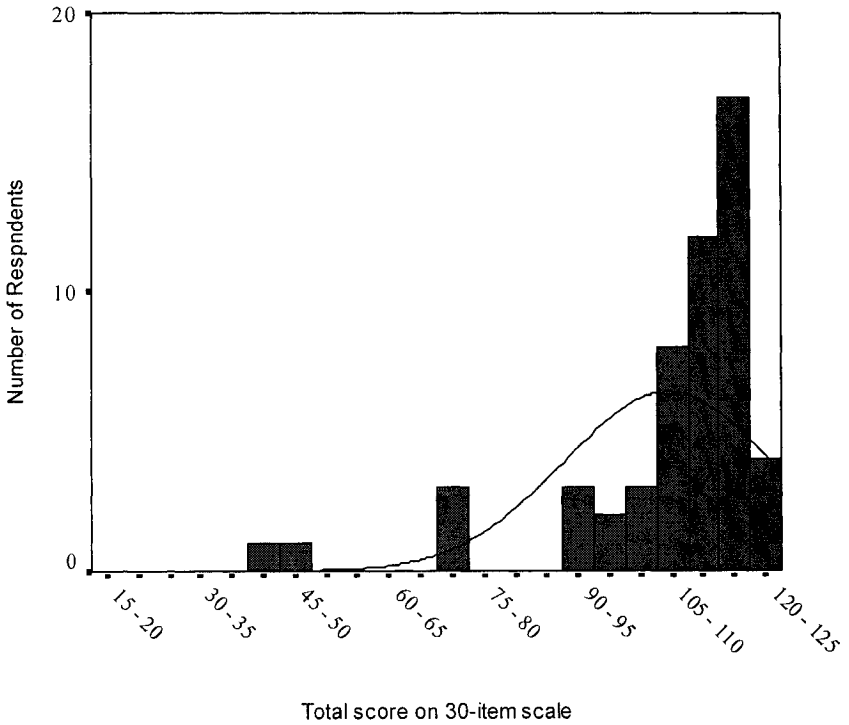


Figure 3: Distribution of Initial Assessment scores of the 30-Item ASK

A scatter plot (Figure 4) comparing the LEALS – SR with the ASK showed a moderate correlation. Figure 4 shows that there is considerable variation in LEALS-SR scores for those who score >110 on the ASK. When the total score for each scale was transformed to a common denominator (100), the mean for the LEALS was found to be 73, while the mean ASK score was 89.4.

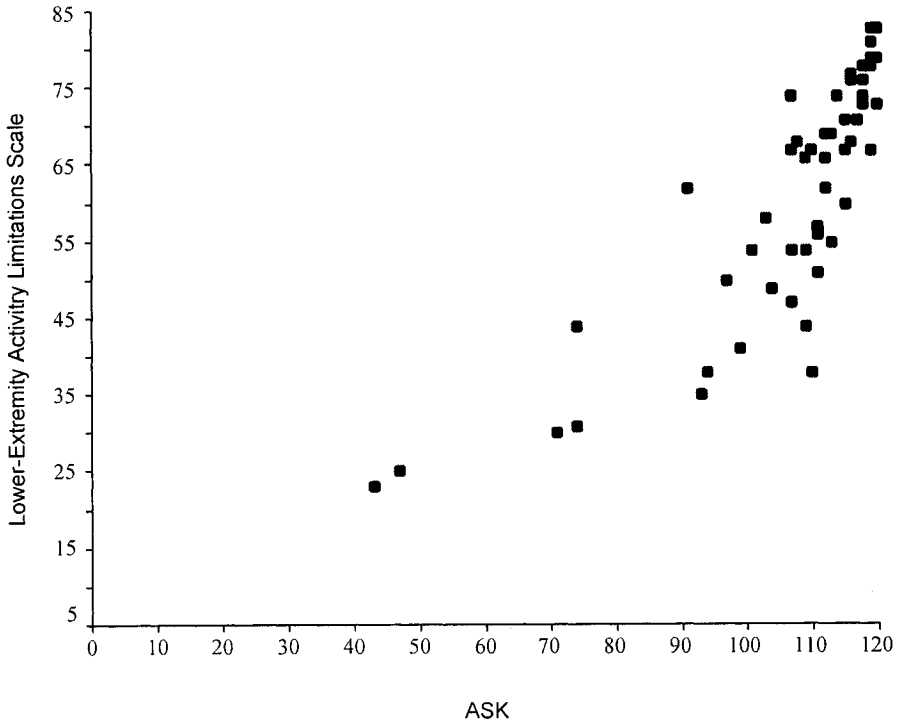


Figure 4: Scatter plot of LEALS- SR and ASK scores at Initial Assessment

Reliability

The internal consistency of both the self-report questionnaire and the functional performance test of the LEALS were found to be good. Cronbach's alpha was 0.95 for the self-report questionnaire and 0.92 for the functional performance test. Individual item to total correlations ranged from 0.41 to 0.80 for the self-report questionnaire and between 0.48 and 0.82 for the functional performance test. Two items in the self-report questionnaire and one item in the functional performance test were found to have item to total correlations of less than 0.5.

Factor Analysis

Principal components analysis (PCA) set to extract eigenvalues greater than one resulted in the extraction of five factors when the self-report items and functional performance test items were entered simultaneously. PCA with an oblimin rotation extracted the same number of factors, but the distribution of items was not as good as the initial solution (Appendix 9). The results of the principal components analysis are shown in Table 12. From the item loadings, two distinct factors were evident: items from the self-report scale formed the first factor while performance items formed a second distinct factor. The scree plot, Figure 5, also provided evidence for a two factor solution. The eigenvalues from factors 3 to 5 lie more or less on a horizontal line suggesting that no significant information is gained from the last three factors.

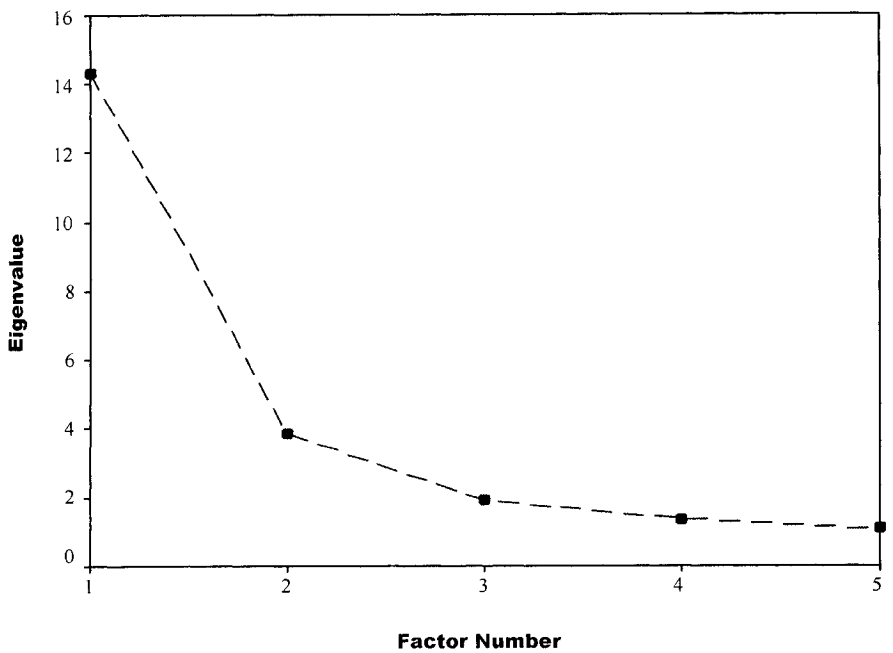


Figure 5: Scree plot of successive eigenvalues following PCA

Table 12: Unrotated factor loadings of items on the LEALS after a PCA

	Component				
	1	2	3	4	5
1. Walking around inside the house	**.795**	-.330	.229	<0.10	<0.10
2. Putting on my socks or shoes	**.569**	**-.400**	.248	.121	**-.474**
3. Squatting	**.635**	<0.10	<0.10	<0.10	.394
4. Lifting an object, like my backpack off the floor	**.476**	-.377	<0.10	.394	<0.10
5. Doing my usual jobs or chores at home	**.754**	-.171	.207	.182	<0.10
6. Getting into or out of a car	**.735**	-.260	.272	.267	-.239
7. Getting into or out of a bus	**.772**	-.302	.259	.238	-.148
8. Walking 2 blocks	**.757**	-.290	<0.10	-.308	<0.10
9. Walking for 30 minutes	**.790**	-.231	<0.10	-.292	<0.10
10. Going up or down a flight of stairs	**.697**	**-.442**	<0.10	-.230	.129
11. Climbing several flights of stairs	**.729**	-.388	<0.10	-.259	.101
12. Standing for 1 hour	**.770**	-.250	<0.10	-.231	.104
13. Sitting for 1 hour	**.485**	**-.450**	.162	.284	.365
14. Running on even ground	**.791**	.193	**-.462**	<0.10	<0.10
15. Running on uneven ground	**.808**	<0.10	**-.458**	<0.10	<0.10
16. Making sharp turns while running fast	**.757**	<0.10	**-.481**	<0.10	<0.10
17. Walking on rough or slippery surfaces	**.701**	-.150	-.204	.215	.114
18. Hopping	**.791**	<0.10	**-.401**	<0.10	<0.10
19. Keeping up with my friends at school or at home	**.738**	-.141	<0.10	-.312	-.114
20. Any of my usual school activities not including gym	**.741**	-.244	.113	<0.10	-.151
21. My usual recreational or sporting activities including gym class	**.810**	.117	-.267	<0.10	<0.10
1. Deep knee bend	.302	**.550**	<0.10	**.483**	.223
2. Repeated sit to stand	.382	**.563**	.375	-.197	<0.10
3. Vertical jump	.384	**.639**	.266	-.152	<0.10
4. Step up	**.545**	**.609**	<0.10	<0.10	-.345
5. Standing hop	**.448**	**.703**	.219	-.184	.154
6. 12 meter hop	**.557**	**.610**	-.305	.110	-.284
7. Balance	.345	**.562**	.370	.134	.193
8. Proprioception	.391	**.549**	.273	-.194	<0.10
9. 22.5 meter walk	**.450**	**.651**	.259	<0.10	.182
10. 22.5-meter run	**.516**	**.624**	<0.10	.175	<0.10

Based on this portion of the analysis it was determined that LEALS scores should be presented as a two-score profile with one score for self-reported function and a separate score for the functional performance test.

Item analyses were conducted by entering self-report and functional performance test items separately into a PCA. Principal components analysis for the functional performance test items resulted in the extraction of a single factor, shown in Table 13. Strong item loadings (>0.5) were established for each item, suggesting that all items support the underlying construct of lower-extremity function.

Principal components analysis for the self-report items also resulted in the extraction of a strong main factor plus two secondary factors, shown in Table 14. On examination of the individual item loadings, all of the self-report items were found to load well on the first factor. In addition, items were found to have lower loadings on the secondary factors. Lastly, the ratio between the first and second eigenvalues was found to be 6:1, suggesting a strong single factor solution. Based on the strong item loadings for the first factor, all of the self-report items were deemed to reflect the same underlying construct.

Table 13: Factor loadings of LEALS- FP items after a PCA

	Factor 1
1. Deep knee bend	.607
2. Repeated sit to stand	.736
3. Vertical jump	.796
4. Standing hop	.803
5. Step up	.856
6. 12 meter hop	.767
7. Balance	.695
8. Proprioception	.686
9. 22.5 meter walk	.832
10. 22.5 meter run	.833

Table 14: Unrotated factor loadings of LEALS- SR items after a PCA

	Factor		
	1	2	3
1. Walking around inside the house	**.838**	.313	<0.10
2. Putting on my socks or shoes	**.635**	**.455**	<0.10
3. Squatting	**.601**	-.105	<0.10
4. Lifting an object, like my backpack off the floor	**.542**	.308	**.527**
5. Doing my usual jobs or chores at home	**.760**	.226	.130
6. Getting into or out of a car	**.754**	.353	.115
7. Getting into or out of a bus	**.807**	.348	.111
8. Walking 2 blocks	**.800**	<0.10	**-.402**
9. Walking for 30 minutes	**.826**	<0.10	-.297
10. Going up or down a flight of stairs	**.785**	.107	-.139
11. Climbing several flights of stairs	**.801**	<0.10	-.203
12. Standing for 1 hour	**.812**	<0.10	-.216
13. Sitting for 1 hour	**.564**	.335	.356
14. Running on even ground	**.747**	**-.564**	.125
15. Running on uneven ground	**.796**	**-.503**	<0.10
16. Making sharp turns while running fast	**.739**	**-.500**	.195
17. Walking on rough or slippery surfaces	**.730**	-.123	.254
18. Hopping	**.783**	-.383	.143
19. Keeping up with my friends at school or at home	**.751**	<0.10	-.378
20. Any of my usual school activities not including gym	**.777**	.127	-.131
21. My usual recreatinal or sporting activities including gym class	**.777**	-.365	<0.10

Item Response Theory

The analysis of item characteristics documented acceptable and stable fit statistics for all items, confirming that all items in the LEALS-SR measured a single construct. The analysis also identified a hierarchical structure among the items. The difficulty level of each item relative to all others is expressed in Table 15 as a range of logit values. For the LEALS-SR, negative logit values indicate easy items and positive logit values indicate difficult items. Item characteristic curves produced for each item (Appendix 10) identified two items with low discriminative ability. Items 3 and 13 (Figures 6, 7) were found to have item discrimination index values of less than two indicating that these

items are unable to discriminate well between subjects with varying levels of lower-extremity function.

Table 15: Marginal maximum likelihood item parameter estimates for LEALS – SR

Item	Item Discrimination Index (a)	Item Difficulty Index (b)
1. Walking around inside the house	6.19	-2.67 to -0.02
2. Putting on my socks or shoes	2.25	-4.53 to -0.38
3. Squatting	1.76	-2.16 to 0.87
4. Lifting an object, like my backpack off the floor	2.77	-3.95 to -0.73
5. Doing my usual jobs or chores at home	3.09	-3.59 to -0.01
6. Getting into or out of a car	3.63	-3.29 to -0.22
7. Getting into or out of a bus	4.87	-1.38 to -0.10
8. Walking 2 blocks	3.33	-1.38 to 0.44
9. Walking for 30 minutes	3.87	-1.02 to 0.79
10. Going up or down a flight of stairs	3.33	-1.39 to 0.61
11. Climbing several flights of stairs	3.23	-0.96 to 1.11
12. Standing for 1 hour	3.21	-1.10 to 0.83
13. Sitting for 1 hour	1.77	-5.46 to -0.21
14. Running on even ground	3.51	-0.66 to 0.92
15. Running on uneven ground	4.24	-0.47 to 1.35
16. Making sharp turns while running fast	2.73	-0.78 to 1.45
17. Walking on rough or slippery surfaces	2.27	-2.13 to 0.92
18. Hopping	2.79	-0.90 to 1.13
19. Keeping up with my friends at school or at home	3.04	-1.32 to 0.44
20. Any of my usual school activities not including gym	3.63	-1.16 to 0.27
21. My usual recreational or sporting activities including gym class	3.27	-0.66 to 1.19

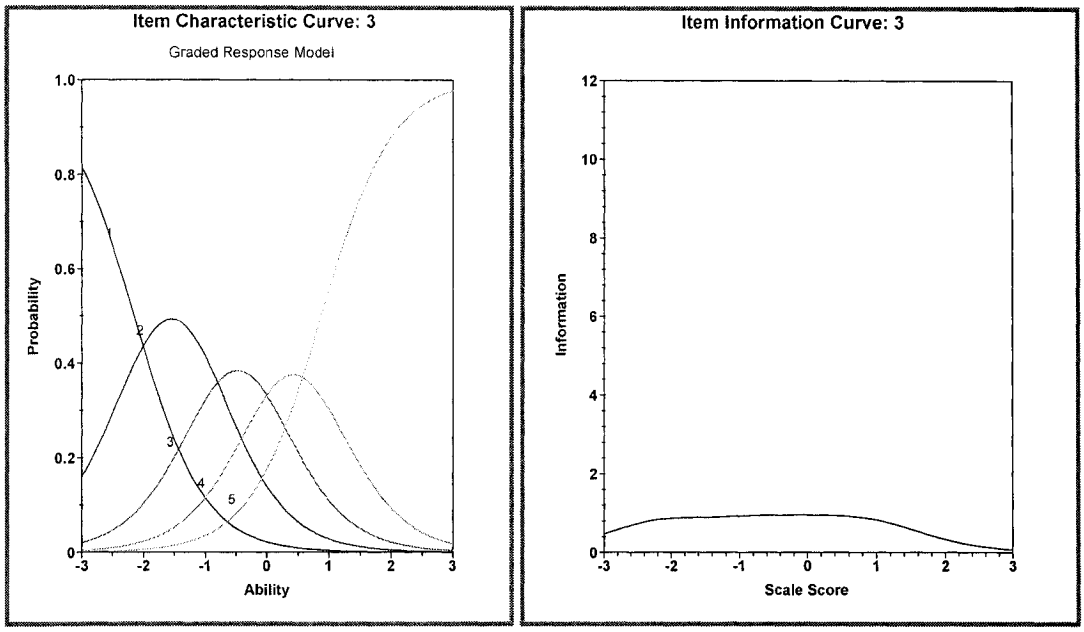


Fig. 6: Item characteristic curves (with five graded response categories) for Item 3 of LEALS-SR, depicting item discrimination and difficulty

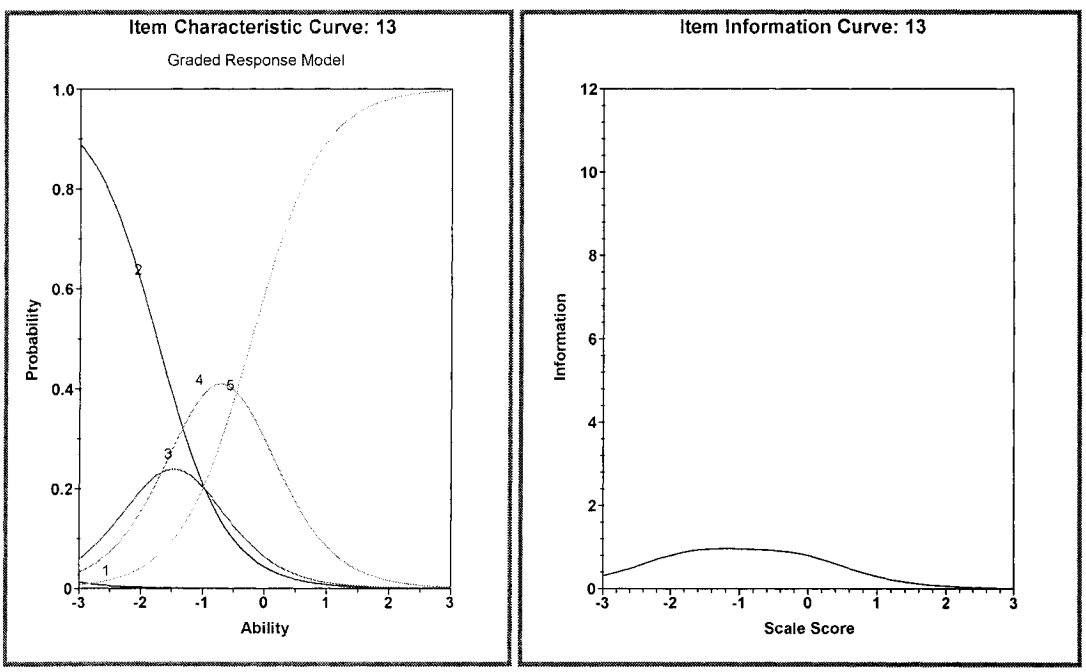


Fig. 7: Item response characteristic curves (with five graded response categories) for Item 13 of LEALS-SR, depicting item discrimination and difficulty

Total test information presented in Fig. 8 suggests that the LEALS-SR is a good measure of lower-extremity function. The test provides the greatest information with the least standard error within the mid range of the construct, lower-extremity function (-2 to 1.5). Increased standard error values are evident at the outer limits; however, these values are still within an acceptable range. Overall test reliability was excellent with alpha = 0.95.

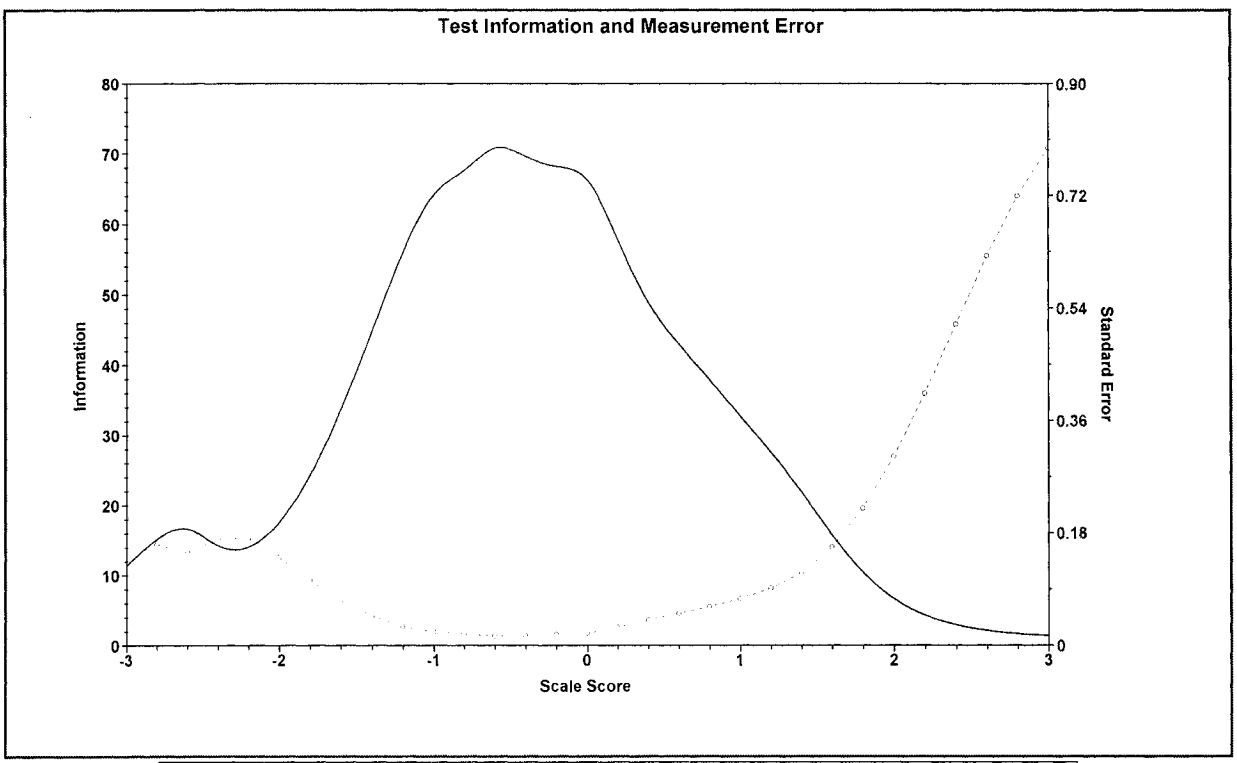


Fig. 8: Total test information and associated standard error curves for the LEALS-SR

Scale Refinement

In order to determine which of the items should be retained in the final scale, item utility analysis was performed. By combining information on response pattern, the items' proximity to the underlying construct, their levels of difficulty and discriminative ability, each item was examined and a decision was made on its inclusion in the final version of the scale. Details of this process are presented in Tables 16 and 17.

In all four items were eliminated from the LEALS-SR as a result of item-utility analysis. Most items were rejected on the basis of several judgment criteria. Item 4 was rejected due to its poor response pattern, low loading on the general lower-extremity function factor and questionable effect on the overall reliability of the scale. Item 3 was rejected primarily due to its inadequate discriminative ability. However, it was also found to have a low item-total correlation. Item 2 was rejected based on its low discriminative ability and highly skewed response pattern. Lastly, Item 13 was rejected for its low discriminative ability and low loading on the general lower extremity function factor.

Scale refinement resulted in the production of a 17-item self-report scale, which was used for the remaining validity analyses.

No items were eliminated from the LEALS-FP following item-utility analysis. However, Item 1 showed borderline results for its measurement of the underlying construct. The item-total correlation was marginal (0.49), the strength of the factor loading was moderate at 0.607 and the items effect on the overall reliability of the scale is questionable.

Table 16: Item Utility Analysis, LEALS – SR

ITEMS	RESPONSE PATTERN	ITEM-TOTAL CORRELATION	CHANGE IN α IF ITEM REMOVED	DISCRIMINATIVE ABILITY (A)	DIFFICULTY (B)
1. Walking around inside the house	62%	.80	-0.004	6.19	-2.67 to -0.02
2. Putting on my socks or shoes	77%	.53	-0.009	2.25	-4.53 to -0.38
3. Squatting	29%	.49	0.001	1.76	-2.16 to 0.87
4. Lifting an object, like my backpack off the floor	89%	.41	0.000	2.77	-3.95 to -0.73
5. Doing my usual jobs or chores at home	64%	.75	-0.003	3.09	-3.59 to -0.01
6. Getting into or out of a car	71%	.67	-0.002	3.63	-3.29 to -0.22
7. Getting into or out of a bus	69%	.77	-0.004	4.87	-1.38 to -0.10
8. Walking 2 blocks	46%	.75	-0.004	3.33	-1.38 to 0.44
9. Walking for 30 minutes	35%	.70	-0.003	3.87	-1.02 to 0.79
10. Going up or down a flight of stairs	40%	.70	-0.003	3.33	-1.39 to 0.61
11. Climbing several flights of stairs	37%	.68	-0.002	3.23	-0.96 to 1.11
12. Standing for 1 hour	33%	.66	-0.002	3.21	-1.10 to 0.83
13. Sitting for 1 hour	65%	.54	-0.000	1.77	-5.46 to -0.21
14. Running on even ground	39%	.69	-0.003	3.51	-0.66 to 0.92
15. Running on uneven ground	31%	.75	-0.004	4.24	-0.47 to 1.35
16. Making sharp turns while running fast	31%	.74	-0.003	2.73	-0.78 to 1.45
17. Walking on rough or slippery surfaces	33%	.68	-0.003	2.27	-2.13 to 0.92
18. Hopping	31%	.76	-0.004	2.79	-0.90 to 1.13
19. Keeping up with my friends at school	46%	.75	-0.004	3.04	-1.32 to 0.44

or at home					
20. Any of my usual school activities not including gym	**50%**	**.74**	**-0.003**	**3.63**	**-1.16 to 0.27**
21. My usual recreational or sporting activities including gym class	**31%**	**.69**	**-0.003**	**3.27**	**-0.66 to 1.19**

Table 17: Item Utility Analysis, LEALS-FP

ITEMS	FACE	CONTENT	ITEM-TOTAL CORRELATION	CHANGE IN α IF ITEM REMOVED	FACTOR LOADINGS
1. Deep knee bend	√	√	**.49**	**0.003**	**.607**
2. Repeated sit to stand	√	√	**.66**	**-0.003**	**.736**
3. Vertical jump	√	√	**.78**	**-0.013**	**.796**
4. Step up	√	√	**.80**	**-0.014**	**.803**
5. Standing hop	√	√	**.80**	**-0.014**	**.856**
6. 12 meter hop	√	√	**.78**	**-0.012**	**.767**
7. Balance	√	√	**.75**	**-0.011**	**.695**
8. Proprioception	√	√	**.61**	**0.002**	**.686**
9. 22.5-meter walk	√	√	**.80**	**-0.007**	**.832**
10. 22.5-meter run	√	√	**.82**	**-0.015**	**.833**

Convergent Validity

Four hypotheses were developed prior to data collection regarding the convergent validity of the new scale. These analyses were performed with the newly abbreviated LEALS- SR and FP scales.

1. **A moderate to high correlation ($r > 0.6$) would exist between self-report LEALS scores and ASK scores.**

As expected, a strong correlation was found between the LEALS – SR scale and the ASK. Correlation coefficients of 0.80 and 0.83 were found respectively for initial assessment data ($n = 52$) and combined data from all the time points ($n = 120$). The relationship was significant to $p < 0.01$.

2. **A low to moderate correlation would exist between lower-extremity function and measures of impairment (ROM, strength, flexibility and pain).**

When self-report and performance scores were compared to impairment measures, the correlations varied greatly. ROM and strength scores were found to have moderate to strong correlations with the performance scale, while pain was found to have a strong correlation with self-reported functional ability. Quality of life was found to correlate poorly with both self-reported and objective measures of functional ability. A summary of the impairment correlation test results can be found in Table 18 below.

Table 18: Correlation Testing of Lower-Extremity Activity Limitations Scale Scores

Variable	LEALS - SR	LEALS - FP
Pain (n = 111)	-0.70**	-0.21*
Range of Motion (n = 112)	0.41**	0.61**
Strength (n = 114)	0.46**	0.66**
Quality of Life$^{\Delta}$ (n = 117)	-0.22**	-0.23*

Pearson's correlation coefficient (2-tailed), ** correlation is significant at the 0.01 level (2-tailed), * correlation is significant at the 0.05 level, $^{\Delta}$ summary score of three questions combined

3. A moderate correlation would exist between the self-report and performance scales of the Lower-Extremity Activity Limitations Scale.

As predicted, LEALS-SR scores were correlated with LEALS- FP scores. The self-report subscale was found to have a moderate correlation (0.57) with the performance subscale with data at the initial assessment. The result remained stable when repeated with the larger sample. This correlation was significant at $p < 0.01$.

4. Patients with more severe conditions would demonstrate lower scores on the self-report and performance scales than patients with mild conditions.

As expected, significant differences in scores on the self-report and performance scales were found when patients were classified by the severity of their condition (Table 19). Patients who were classified as having a condition which would severely affect their functional ability scored lower on both the self-report and functional performance scales. Patients with a severe condition were found to have a mean score of 50 on the performance scale, and 41 on the self-report scale. Children with mild or moderate conditions obtained mean scores of 57 and 75 on the performance scale and 49 and 57 on the self-report scale, respectively. The differences in scores across severity were

significant, $p < 0.01$, for both subscales with initial assessment data and repeated measures data.

Table 19: Mean scores on the LEALS – SR and FP by severity of condition

	LEALS - SR		LEALS - FP	
Severe				
Mean (SD)	35.3 (±11.3)	40.6 (±12.7)	40.0 (±23.4)	49.5 (±24.5)
N	9	13	9	13
Moderate				
Mean (SD)	46.6 (±13.5)	49.0 (±14.2)	55.3 (±19.0)	57.0 (±17.0)
N	19	32	20	33
Mild				
Mean (SD)	52.0 (±12.4)	57.0 (±11.0)	73.4 (±7.5)	75.0 (±8.3)
N	22	54	22	54
Total assessments	50	99	51	100

Sensitivity to Change

Sensitivity to change was examined by comparing the mean change per day for the LEALS-SR and the ASK. Scale scores were converted to a score between 0 and 100 for both scales to allow for direct comparisons. Analyses were performed on a subsample of children who were deemed to have changed based on patient and clinician prognostic ratings. Overall the LEALS-SR was found to be more sensitive to change than the ASK. Between baseline and three weeks follow-up, the mean change per day was found to be 0.65 and 0.19 for the LEALS-SR and the ASK respectively (Figure 9).

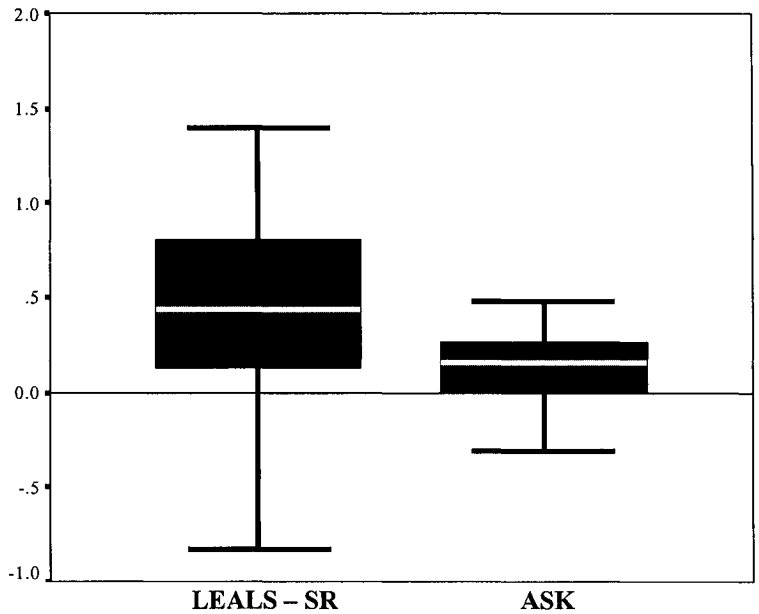


Figure 9: Mean change scores for the LEALS – SR and the ASK at 3 weeks follow-up (n = 40)

At 6 weeks follow-up the mean change per day for LEALS- SR was 0.42 while the mean change on the ASK was 0.24 (Figure 10).

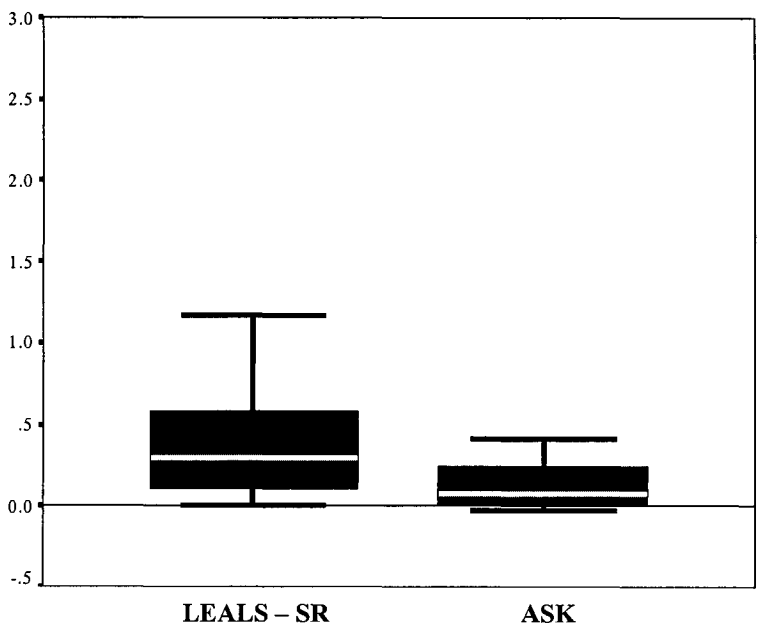


Figure 10: Mean change scores for the LEALS-SR and the ASK at 6 weeks follow-up (n = 26)

When sensitivity to change scores for the LEALS-SR and ASK were compared across severity levels, the LEALS-SR was found to be more sensitive to change for all levels of severity (Figure 11).

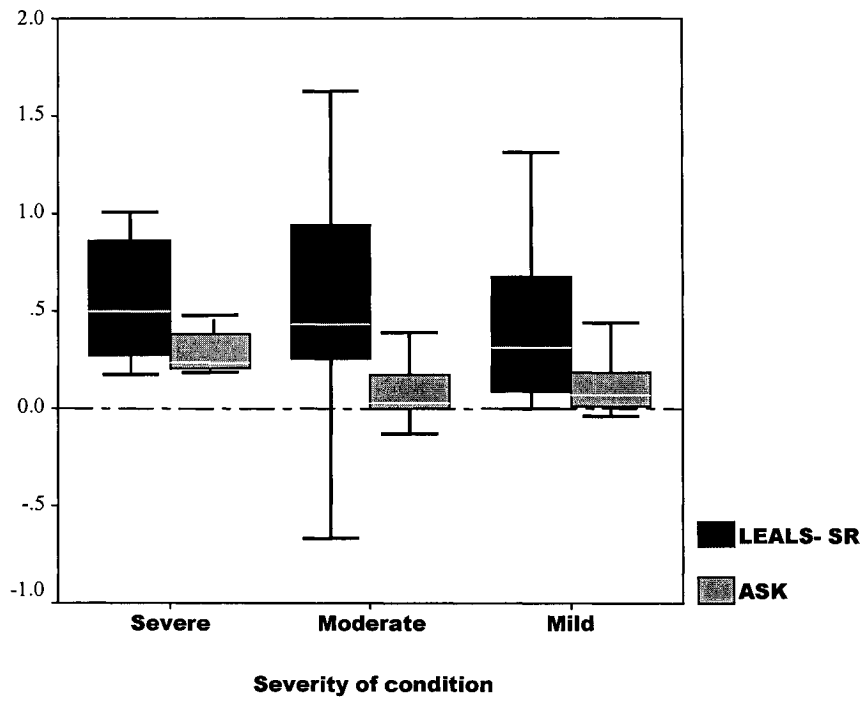


Figure 11: Mean change scores for the LEALS-SR and the ASK at 3 weeks follow-up, by severity of condition (n = 39)

Regression Analysis

Prior to conducting the regression analysis, descriptive statistics were produced for the dependent and independent variables (Table 20). A correlation matrix was produced to examine the relationship amongst variables and scatter plots were produced to examine the relationship of each independent variable with the dependent variable (Appendix 11). Following basic descriptive statistics, two multiple linear regressions were performed; one with LEALS –SR scores as the dependent variable and one with LEALS – FP scores

as the dependent variable. Using stepwise selection with a 0.05 significance level for entry into the model, three variables were identified as being significant predictors of LEALS - SR and FP test scores (Table 21).

The linear regression for LEALS – SR scores suggests that pain level, strength and the patient's response to their condition are significant factors for predicting LEALS – SR scores. The slope of the line for pain suggests that each unit increase in pain (10 cm scale) results in a decrease in total score on the LEALS- SR by 2.5 on average. The slope of the line for strength suggests that each unit increase in overall strength results in an increase of 2.0 on the LEALS – SR total score. Lastly, the slope of the line for the response VAS suggests that more exaggerated responses to the condition decrease the overall LEALS – SR score. The regression model provides a good fit to the observed data. Seventy-three percent (adjusted R^2) of the variability in LEALS – SR scores can be explained by the linear relationship with the three independent variables.

From the linear regression on LEALS- FP scores, three variables were identified as being significant: range of motion, anticipated impact of the injury on ADLs and pain. The slope of the line for range of motion suggests that each unit increase in range of motion results in an increase in total score on the LEALS- FP by 0.5 on average. The slope of the line for impact of the injury suggests that children with an injury that is anticipated to have a minimal impact on function score higher on the LEALS – FP than those with an injury anticipated to have a greater impact on function. Lastly, the slope of the line for pain suggests that each unit increase in pain results in a decrease in LEALS – FP scores by 1.5 on average. The regression model provides a good fit to the observed

data: 52% (adjusted R^2) of the variability in LEALS – FP scores can be explained by the linear relationship with the three independent variables.

Table 20: Descriptive statistics for the independent and dependent variables used in multiple linear regression analyses (n = 120)

	Mean	Sd	Skewness	Kurtosis
LEALS -SR	47.58	13.77	-.63	.16
LEALS - FP	66.79	17.17	-1.47	2.60
Strength	18.45	2.36	-2.89	12.52
Range of motion	-4.70	14.83	-4.55	30.76
Pain	3.00	2.55	.55	-.78
Quality of Life*	3.12	2.31	.56	-.56
Impact VAS	6.60	2.48	-.38	-1.07
Response VAS	5.49	1.24	1.07	2.31

*summary score of three questions combined

Table 21: Multiple linear regression coefficients with associated standard error and significance levels for the LEALS- SR and FP

Model	B	Std. Error	t	Sig.	R^2	Adjusted R^2
LEALS - SR					0.745	**0.726**
Constant	43.87	8.98	4.885	.000		
Pain	-2.517	.469	-5.367	.000		
Strength	1.976	.413	4.780	.000		
Response VAS	-3.768	.884	-4.265	.000		
LEALS – FP					0.536	**0.520**
Constant	58.07	6.21	9.349	.000		
Range of motion	.499	.099	5.065	.000		
Impact VAS	2.446	.817	2.996	.005		
Pain	-1.561	.708	-2.203	.033		

DISCUSSION

Summary of Findings

During phase I, a two part scale was developed measuring lower-extremity function in adolescents. The subsequent administration of the scale to an adolescent population decreased the number of items in the scale and presented a preliminary assessment of the scale's reliability, validity and sensitivity to change.

The original conceptual framework was developed by the primary investigator and a panel of experts and was based in part on the ICIDH-2. Reviewing existing questionnaires and surveying clinical experts led to the development of 48 self-report items and 26 functional performance test items hypothesized to measure lower-extremity function. The 48 self-report items were reduced to 26 through a consensus process involving all investigators. The same process resulted in the removal of 13 items from the functional performance test. The 26 self-report items were further reduced to 21 while the 13 functional performance test items were reduced to ten following a review for content validity.

To further evaluate the items relationship to the conceptual framework, a factor analysis was undertaken using the responses from 122 repeated measures assessments. The factor analysis supported the two hypothesized dimensions of lower-extremity function: self-report and performance; this provided evidence that the results of the LEALS should be presented as a profile with two component scores.

Additional analyses were performed to evaluate the utility and validity of each item in the two subscales. From this process, the LEALS-SR was reduced from 21 to 17 items. One additional item in the LEALS - SR produced questionable results. Item 17, walking

on rough or slippery surfaces, produced borderline results for its discriminative ability however, performed well on the remaining item utility analyses. Item 17 was retained in the final scale for further consideration but may benefit from rewording to improve its clarity.

No items were removed from the LEALS-FP as a result of the item utility analyses. However, Item 1 produced lacklustre results. These results may be due in part to the limited sample. Item 1 was designed to measure at the low level of lower-extremity function but our sample included very few low functioning adolescents. It has been retained for further consideration but may need to be eliminated in the future.

Following item utility analyses, the remaining 17 self-report items and ten functional performance test items were used to examine the validity, reliability and sensitivity to change. On the basis of these results, the LEALS displayed good content and construct validity, excellent reliability and adequate responsiveness.

The content validity of the LEALS was supported by the procedures used in the item development phase, including reviews of existing instruments and review by an expert panel. It was also supported by the LEALS score distributions which showed adequate range and variability with no ceiling or floor effects. IRT analysis also confirmed that self-report test items were distributed adequately across the continuum of disability.

Internal consistency was within reasonable limits for survey development[78]. The values for Cronbach's alpha suggested that items in both the self-report scale and functional performance test were representative of the construct being studied. Similarly, the high Cronbach's alpha values (LEALS- SR = 0.95; LEALS - FP = 0.92) for both scales suggested that the subscales contribute to a cohesive and consistent health status

profile. Excellent inter-rater reliability was established for the functional performance test.

Construct validity was supported by the convergent analyses. All four of the predicted relationships were supported by the results of this study. As expected, lower-extremity function as measured by the LEALS –SR was strongly correlated with ASK scores. The correlation was found to be high but was limited by the ceiling effects demonstrated in the ASK. Similarly, self-report scores were moderately correlated to functional performance test scores. This finding suggests that the LEALS may be particularly useful in helping to form a more global clinical picture than a scale which only evaluates perceived function or observed function. As in previous research[3, 7, 12, 13], clinical measures (e.g., ROM, strength, pain) were only moderately or weakly associated with the construct. These clinical measures are meaningful indicators of impairment for clinicians but may be less meaningful indicators of functional ability for patients. The LEALS may be better able to identify the adaptations that patients make to allow them to function adequately in day-to-day activities (e.g., adjusting gait to accommodate the injured joint), despite limitations in ROM or strength. This result confirms the importance of using multiple outcome measures in order to obtain a global clinical picture.

Lastly, LEALS scores differed significantly when patients were classified by the severity of their condition. Patients who were classified, by their therapist, as having a condition that would severely affect their functional ability scored lower on both the self-report and functional performance scales. This result is crucial as it provides additional evidence to support the underlying construct. If patients did not differ significantly when

classified by the severity of their condition then the ability of the LEALS to measure function would be questionable.

The responsiveness of the LEALS was substantiated by the moderate to high correlations with patient and therapist reported change scores. In addition, the capacity of the LEALS- SR to detect change in lower-extremity function was found to be superior to that of the ASK for this specific patient population. This result is likely due to the ceiling effects which were reported in the ASK. In the case of a ceiling effect, there is restricted range for improvement because patients begin at the high level of function on the scale. The implication of ceiling effects is to lower the capacity to detect clinically important change in all patients. Therefore, the LEALS superior ability to detect change in the population of interest was not surprising.

Adoption of LEALS into clinical practice

The LEALS was designed to measure one aspect of the ICIDH-2 framework, activity limitation therefore it should not be used assess impairment or restrictions in participation in the community. However, it provides a complementary perspective on the disability process and should be used in conjunction with other measures to generate a global clinical picture.

It is easy to administer and potentially applicable for a wide-range of disability levels and orthopedic conditions in the lower-extremities. The LEALS can be used by clinicians to establish initial levels of lower-extremity function, track progress, or set goals. The LEALS-SR takes minimal time (ten minutes) to administer and is easy to score. It requires no training and can be completed independently by most adolescents ages ten

and over. In order to administer the LEALS- FP, some training is required. Therapists should be familiar with the test items and the standard instructions for patients. However, administration is straightforward and does not require complicated equipment or a lot of time. In order to calculate an overall score for the LEALS-FP a conversion chart is required. The chart is needed to transform individual item scores into a common denominator resultant in equal weightings for each item. This constitutes an extra step for therapists but does not require difficult calculations and will result in a total score for the subscale that can be compared over time.

Limitations of the Study

It is very difficult to develop a valid, reliable and clinically relevant measurement instrument. Many of the limitations of this project are due to time constraints. Ideally, the study would have included a larger and more diverse sample. The LEALS was conceived as a measure applicable to a broad spectrum of lower-extremity problems but our sample did not include any patients with hip conditions. In addition, a limited number of participants presented with conditions thought to severely affect their function. These limitations in sample size and diversity limit the generalizability of the measure. The characteristics of the sample reflect numerous external factors including the length of the wait list and time of the year. The sample while not ideal is reflective of the constraints in current health care as well as seasonal factors. Limited funding can result in longer waiting lists which affect the type of injuries or conditions that are seen. Seasonal factors also have an affect on the type of cases, for example elective surgeries for children are

more likely to be scheduled during the summer. Both of these factors impacted the characteristics of the sample recruited.

The results on the reliability of the LEALS were also limited by the sample. Initially we planned to assess the test-retest reliability of the LEALS. However, we were unable to obtain the appropriate sample to carry out this assessment. For test-retest reliability, assessments must be carried out before there is any change in the adolescents' condition. This required re-assessments to be conducted within 72 hours and only on children with a chronic condition. Because of the length of the wait list and compressed work schedule of the therapists, no children with chronic conditions were seen within 72 hours of the initial assessment. Therefore, we are unable to comment on the stability of the LEALS.

Furthermore, the use of repeated-measures data for some analyses is a limitation. Because some data represents information from a single test subject at two or three different time points, the variability of this data is probably less than it would be if separate subjects had been used. Although the point estimates should remain stable, the use of this type of data probably resulted in narrower confidence intervals than would occur in an independent sample.

Although the use of repeated measures data can be considered a limitation, its use was necessary in order to achieve adequate numbers for some of the analyses. It has been suggested that in order to perform factor and regression analyses, there should be at least five times (some authors as many as 20 times) more respondents in the sample than variables to be analyzed [30, 74, 79, 80, 81]. Repeated measures data has also been used in at least one recently published article[77].

In order to avoid the underestimation in variation resulting from the use of repeated measures data, a larger independent sample could have been recruited or statistical analyses controlling for the differences in variation could have been performed. For the regression analysis a mixed effects model could have been used to correct the variation estimates. However, the analysis would have been very complicated because the data was unbalanced (i.e. differing numbers of assessments for each participant); to provide an adequate correction a nest factor and a crossover factor would have been required. In addition, the significant variables identified during regression analysis using repeated measures data were the same as those selected during regression with initial assessment data.

In addition to being influenced by the use of repeated measures data, the factor analysis results may have been affected by the type and distribution of the variables included. Although principal components analysis makes no special assumptions about the data being continuous or normally distributed most methods of estimation are based on Pearson's Correlation Coefficient which assumes continuous, normally distributed data. The effect of categorical data on the results is largely unknown but according to Fayers and Machin scales with as few as 5-categories can get stable results[81]. The results of the factor analyses were also supported by the item response analyses which confirmed that the LEALS-SR formed a single construct.

Several other statistical calculations may have had a minor influence on the results. Intraclass correlation coefficients were used to calculate the inter-rater reliability for two discrete items in the LEALS-FP. A weighted Kappa would have been a slightly more accurate method. However, as three examiners were involved weighted Kappa's would

be required between every possible pair of examiners. This would have complicated the interpretation of the results. In addition, the data contained a very large number of response categories which were normally distributed suggesting that the use of an ICC would have a minimal effect on the estimates. Lastly, the standardization method used to calculate sensitivity to change did not consider potential differences in the standard deviations of the two scales being compared. A different approach could have been used to create standardized scores while correcting for differences in standard deviation. However, this would have complicated the calculation for mean change in scale score per day. In addition, the standard deviations were found to be similar suggesting that the more complicated method of standardization might not be required.

Future Directions

Instrument validation is an ongoing process; additional studies with diverse samples and different criteria are needed to verify the preliminary results. These studies should aim to include other populations with limited lower extremity function such as amputees or patients post-surgery. Future samples should also include a greater number of patients with severe or acute conditions. Studies must also be undertaken in other centres, for the original developers of the scale almost always report better validity results than subsequent users.

The validity of the LEALS-SR when completed by a proxy is also of interest. The current version of the scale has been examined in adolescents. However, for younger children it would be interesting to examine the validity of the results when completed by a parent or guardian.

Additional research should also be conducted on the reproducibility of LEALS results. An assessment of test-retest reliability is essential to ensure that results on the LEALS remain stable when no change has occurred in lower-extremity function.

The utility of the remaining items in the LEALS must also be re-assessed. Although the current study used several different methods to assess the usefulness of each item, these results must be verified with different samples.

Lastly, outcome measurement in rehabilitation must continue to keep pace with advances in treatment and the changing perspective on disability. In recent years, outcome measurement in rehabilitation has broadened to include evaluations of activity limitation in addition to measures of impairment. However, we must continue to expand our perspective and develop valid and reliable ways to measure to participation restrictions, the final dimension in the WHO classification of Functioning and Disability.

Conclusion

This study has resulted in the development of an outcome measure evaluating lower-extremity function in adolescents with orthopaedic conditions. It is brief, easy to administer and identifies initial levels of lower extremity function as well as changes in function over time.

By examining the results from the self-report questionnaire and functional performance test, physiotherapists may be better able to identify the factors limiting lower extremity function.

REFERENCES

1. Statistics Canada (2001) Participation and Activity Limitation Survey [Internet]. Available at: www.statcan.ca/Daily/English/030729/d030729b.html

2. Haley SM, Russo KJ. Musculoskeletal outcomes in pediatric physical therapy practice. Orthopedic Physical Therapy 2000; 9: 17-36.

3. Jette AM. State of the art functional status assessment. In: Rothstein JM, Rothstein JMS (eds). Measurement in physical therapy. New York: Churchill Livingstone; 1985: 137-168.

4. Young NL, Wright JG. Measuring pediatric physical function. Journal of Pediatric Orthopedics 1995; 15: 244-253.

5. World Health Organization. ICIDH-2: International Classification of Functioning and Disability. 54th World Health Assembly, Geneva: World Health Organization; 2001.

6. Vitale MG, Levy DE, Johnson MG, Gelijns AC, Moskowitz AJ et al. Assessment of quality of life in adolescent patients with orthopedic problems: are adult measures appropriate. Journal of Pediatric Orthopedics 2001; 21: 622-628.

7. Lee CE, Simmonds MJ, Novy DM, Jones S. Self-reports and clinician measured physical function among patients with low back pain: a comparison. Archives of Physical Medicine and Rehabilitation 2001; 82: 227-231.

8. Clark NC. Functional performance testing following knee ligament injury. Physical Therapy in Sport 2001; 2: 91-105.

9. Hart DL, Velozo CA, Lai JS, Dobrzykowski EA. The reliability and validity of the Orthopedic Rehabilitation Outcome Scale. Journal of Rehabilitation Outcome Measures 1997; 1: 1-7.

10. Novy DM, Simmonds MJ, Lee CE. Physical performance tasks: what are the underlying constructs? Archives of Physical Medicine and Rehabilitation 2002; 83:44-47.

11. Young NL, Williams JI, Yoshida KK, Wright JG. Measurement properties of the Activities Scale for Kids. Journal of Clinical Epidemiology 2000; 53: 125- 137.

12. Salen BA, Spangfort EV, Nygren AL, Nordemar R. The Disability Rating Index: an instrument for the assessment of disability in clinical settings. Journal of Clinical Epidemiology 1994; 47:1423- 1435.

13. Lindstrom I, Ohlund C, Nachemson A. Validity of patient reporting and predictive value of industrial physical work demands. Spine 1994; 19: 888- 893.

14. Riddle DL, Stratford PW, Binkley JM. Sensitivity to change of the Roland-Morris Back Pain Questionnaire: part 2. Physical Therapy 1998; 78: 1197-207.

15. Lacker JM, Carosella AM, Feuerstein M. Pain expectancies, pain, and functional self-efficacy expectancies as determinants of disability in patients with chronic low back disorders. J Consult Clin Psychol 1996; 64: 212-20.

16. Noyes F, Barber S, Mangine R. Abnormal lower limb symmetry determined by function hop tests after anterior cruciate ligament rupture. American Journal of Sports Medicine 1991; 19: 513-518.

17. Wilk K, Romaniello W, Soscia S, Arrigo C, Andrews J. The relationship between subjective knee scores, isokinetic testing, and functional testing in the ACL-reconstructed knee. Journal of Orthopedic and Sports Physical Therapy 1994; 20: 60-73.

18. Purdam CR, Cook JL, Hopper DM, Khan KM, VIS tendon study group. Discriminative ability of functional loading tests for adolescent jumper's knee. Physical Therapy in Sport 2003; 4: 3-9.

19. Hopper DM, Goh SC, Wentworth LA, Chan D, Chau J et al. Test-retest reliability of knee rating scales and functional hop tests one year following anterior cruciate ligament reconstruction. Physical Therapy in Sport 2002; 3: 10-18.

20. Lephart SM, Perrin DH, Fu F, Minger K. Functional performance tests for the anterior cruciate insufficient athlete. Athletic Training 1991; 26: 44-50.

21. Noyes FR, Barber SD, Mooar LA. A rationale for assessing sports activity levels and limitations in knee disorders. Clinical Orthopedics 1989; 246: 238-249.

22. Lysholm J, Gillquist J. Evaluation of knee ligament surgery results with special emphasis on use of a scoring scale. American Journal of Sports Medicine 1982; 10: 150-154.

23. American Academy of Orthopedic Surgeons. Scoring algorithms for the lower limb outcomes data collection instrument, version 2.0. Rosemont, IL: American Academy of Orthopedic Surgeons; 1998.

24. Irrang JJ, Snyder-Mackler L, Wainner RS, Fu FH, Harner CD. Development of a patient-reported measure of function of the knee. Journal of Bone and Joint Surgery (American Version) 1998; 31: 1132-1145.

25. Kujala UM, Jaakkola LH, Koskinen SK, Taimela S, Hurme M et al. Scoring of patellofemoral disorders. Arthroscopy. 1993; 9: 159-163.

26. Chesworth BM, Culham EG, Tata GE, Peat M. Validation of outcome measures in patients with patellofemoral syndrome. Journal of Orthopedic and Sports Physical Therapy 1989; 10: 302-308.

27. Kantz ME, Harris WJ, Levitsky K, Ware JE, Davies AR. Methods for assessing condition-specific and generic functional status outcomes after total knee replacement. Medical Care 1992; 30: 240-252.

28. Tegner Y, Lysholm J. Rating systems in the evaluation of knee ligament injuries. Clinical Orthopedics. 1985; 198: 43-49.

29. Hawker G, Melfi CA, Paul J, Green R, Bombardier C. Comparison of a generic (SF-36) and disease specific (WOMAC) instrument in the measurement of outcomes after knee replacement surgery. Journal of Rheumatology 1995; 22: 1193-1196.

30. McDowell I, Newell C. Measuring Health: A guide to rating scales and questionnaires. Oxford University Press, Oxford, 1996.

31. Wilkin D, Hallam L, Doggett, MA. Measures of need and outcome for primary health care. Oxford University Press, Oxford, 1992.

32. Finch E, Brooks D, Stratford P, Mayo N. Physical Rehabilitation Outcome Measures: A Guide to Enhanced Clinical Decision Making. 2nd Edition. BC Decker, Hamilton, 2002.

33. Vitale MG, Levy DE, Moskowitz AJ, Gelijns AC, Spellmann M et al. Capturing quality of life in pediatric orthopedics: two recent measures compared. Journal of Pediatric Orthopedics 2001; 21:629-635.

34. Wright JG, Young NL. Outcome assessment in children. In: Swiontkowski G (ed). Skeletal trauma in children. Philadelphia: Lippincott-Raven. 1996:149-160.

35. Daltroy LH, Liang MH, Fossel AH, Goldberg MJ. The POSNA pediatric musculoskeletal functional health questionnaire: report on reliability, validity, and sensitivity to change. Journal of Pediatric Orthopedics 1998; 18: 561-571.

36. Feldman AB, Haley SM, Coryell J. Concurrent and construct validity of the Pediatric Evaluation of Disability Inventory. Physical Therapy 1990; 70: 602-610.

37. Kessler S, Pfonder T, Nelitz M, Puhl W, Gunther KP. The Pediatric Musculoskeletal Functional Health Questionnaire. Zeitschrift fur Orthopadie und Ihre Grenzgebiete 2001; 139: 134- 137.

38. Walker LS, Greene JW. The functional disability inventory: measuring a neglected dimension of child health status. Journal of Pediatric Psychology 1991; 16: 39-58.

39. Schunk C, Rutt R. The Therapeutic Associates Outcome System (TAOS) Functional Index. Journal of Rehabilitation Outcomes Measurement 1998; 2: 55-61.

40. Law M, Usher P. Validation of the Klein-Bell Activities of Daily Living Scale for children. Canadian Journal of Occupational Therapy 1988; 55: 63-8.

41. Hayley SM, Russo KJ. Musculoskeletal outcomes in pediatric physical practice. Orthopaedic Clinics of North America 2000; 9: 17-36.

42. Pencharz J, Young NL, Owen JL, Wright JG. Comparison of three outcome instruments in children. Journal of Pediatric Orthopedics 2001; 21: 425-432.

43. Young NL, Yoshida KK, Williams JI, Bombardier C, Wright JG. The role of children in reporting their physical function. Archives of Physical Medicine and Rehabilitation 1995; 76:913-8.

44. Haynes RJ. Sullivan E. The Pediatric Orthopedic Society of North America Functional Health questionnaire: an analysis of normals. Journal of Pediatric Orthopedics 2001; 21: 619-621.

45. Binkley JM, Stratford PW, Lott SA, Riddle DL. The North American Orthopedic Rehabilitation Research Network. The Lower Extremity Functional Scale: Scale development, measurement properties, and clinical application. Physical Therapy 1999; 79: 371-383.

46. Tabor MA, Davies GJ, Kernozek TW, Negrete RJ, Hudson V. A multicenter study of the test-retest reliability of the Lower Extremity Functional Test. Journal of Sport Rehabilitation 2002; 11:190-201.

47. Delitto A, Irrang J, Harner C, Fu F, Nessi S. Relationship of isokinetic quadriceps peak torque and work to one legged hop and vertical jump in ACL reconstructed subjects [abstract] Physical Therapy 1993; 73.

48. Eastlack M, Axe M, Snyder-Mackler L. Laxity, instability, and functional outcome after ACL injury: copers versus noncopers. Medicine and Science in Sports and Exercise 1999; 31: 210-215.

49. Lephart S, Perrin D, Fu F, Gieck J, McCue F, Irrang J. Relationship between selected physical characteristics and functional capacity in the anterior cruciate ligament insufficient athlete. Journal of Orthopedic and Sports Physical therapy 1992; 16: 174-181.

50. Lephart S, Perrin D, Fu F, Minger K. Functional performance tests for the anterior cruciate ligament insufficient athlete [abstract]. Journal of Athletic Training 1991; 26: 44-

51. Ostenberg A, Roos E, Ekdahl C, Roos H. Isokinetic knee extensor strength and functional performance in healthy female soccer players. Scandinavian Journal of Medicine and Science in Sports 1998; 8: 257-264.

52. Petching R, Baron R, Albrecht M. The relationship between isokinetic quadriceps strength test and hop tests for distance and one-legged vertical jump test following anterior cruciate ligament reconstruction. Journal of Orthopedic and Sports Physical Therapy 1998; 28: 23-31.

53. Blackburn J, Morrissey M. The relationship between open and closed kinetic chain strength of the lower limb and jumping performance. Journal of Orthopedic and Sports Physical Therapy 1998; 27: 430-435.

54. Gauffin H, Ekstrand J, Arnesson L, Tropp H. Vertical jump performance in soccer players: a comparative study of two training regimes. Journal of Human Movement Studies 1989; 16: 215-224.

55. Greenberger H, Paterno M. Comparison of an isokinetic strength test and a functional performance test in the assessment of lower extremity function [abstract]. Journal of Orthopedic and Sports Physical Therapy 1994; 19: 61.

56. Kramer W, Triplett N, Fry A, Koziris L, Bauer J et al. An in-depth sports medicine profile of women college tennis players. Journal of Sport Rehabilitation 1995; 4: 79-98.

57. Ostenberg A, Roos, Ekdahl C, Roos H. Isokinetic knee extensor strength and functional performance in healthy female soccer players. Scandinavian Journal of Medicine and Science in Sports 1998; 8: 513-518.

58. Pincivero D, Lephart S, Karunakara R. Relation between open and closed kinetic chain assessment of knee strength and functional performance. Clinical Journal of Sport Medicine 1997; 7: 11-16.

59. Bombardier C, Tugwell P. Methodological considerations in functional assessment. Journal of Rheumatology 1987; 14 Suppl: 6-10.

60. Streiner DL, Norman GR. Health Measurement Scales. 2nd ed. Oxford University Press. Toronto. 1995.

61. Nuckols RC. A note on pre-testing public opinion questions. Journal of Applied Psychology 1953; 37:119- 120.

62. Oleckno W. Essential Epideiology. Illinois: Waveland Press, 2002.

63. Traub R., Rowley GL. Understanding Reliability. Instructional Topics in Educational Measurement. National Council on Measurement in Education. Spring 1991.

64. Jenkinson C, Mayou R, Day A, Garrat A, Juszczak E. Evaluation of the Dartmouth COOP charts in a large-scale community survey in the United Kingdom. Journal of Public Health Medicine 2002; 24: 106-111.

65. Lowry R. (2004) Vassar College [Internet] Available from: http://faculty.vassar.edu/lowry/corr_stats.html

66. Pencharz J, Young NL, Owen JL, Wright JG. Comparison of three outcomes instruments in children. Journal of Pediatric Orthopedics 2001; 21: 425-432.

67. Caraceni A, Cherny N, Fainsinger R, Kaasa S, Poulain P et al. Pain measurement tools and methods in clinical research in palliative care: recommendations of an expert working group of the European Asscoiation of Palliative Care. Journal of Pain and Symptom Management 2002; 23: 239-255.

68. Wasson JH, Kairys SW, Nelson EC, Kahishman N, Baribeau P. A short survey for assessing health and social problems of adolescents. The Journal of Family Practice 1994; 38: 489-494.

69. Clarkson HM, Gilewich GB. Musculoskeletal assessment – joint range of motion and manual muscle strength. Baltimore: Williams & Wilkins, 1989.

70. Medical Research Council. Aids to the examination of the peripheral nervous system. London: Her Majesty's Stationery Office, 1976: 14.

71. Kendall FP, McCreary EK, Provance PG, Kendall HO. Les muscles: bilan et étude fonctionnels, anomalies et douleurs posturales. 4e ed. Paris: Pradel, 2001.

72. Florence JM, Pandya S, King WM, Robison JD, Baty J et al. Intrarater reliability of manual muscle test (Medical Research Council scale) grades in Duchenne's muscular dystrophy. Physical Therapy 1992; 72: 115-122.

73. Nunnally JC, Bernstein IH. Psychometric Theory. 3rd ed. New York: McGraw-Hill, 1994: 264-5.

74. Stevens J. Applied Multivariate Statistics for the Social Sciences. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates, 1992.

75. Statistical Programs for Social Sciences. SPSS, Version 11.0: Statistics Coach.

76. Safrit MJ, Cohen AS, Costa MG. Item response theory and the measurement of motor behavior. Research Quarterly for Exercise and Sport 1989; 60: 325-335.

77. Van Excel NJA, Scholte op Reimer WJM, Koopmanschap MA. Assessment of post-stroke quality of life in cost-effectiveness studies: The usefulness of the Barthel Index and the EuroQoL-5D. Quality of Life Research 2004; 13: 427-433.

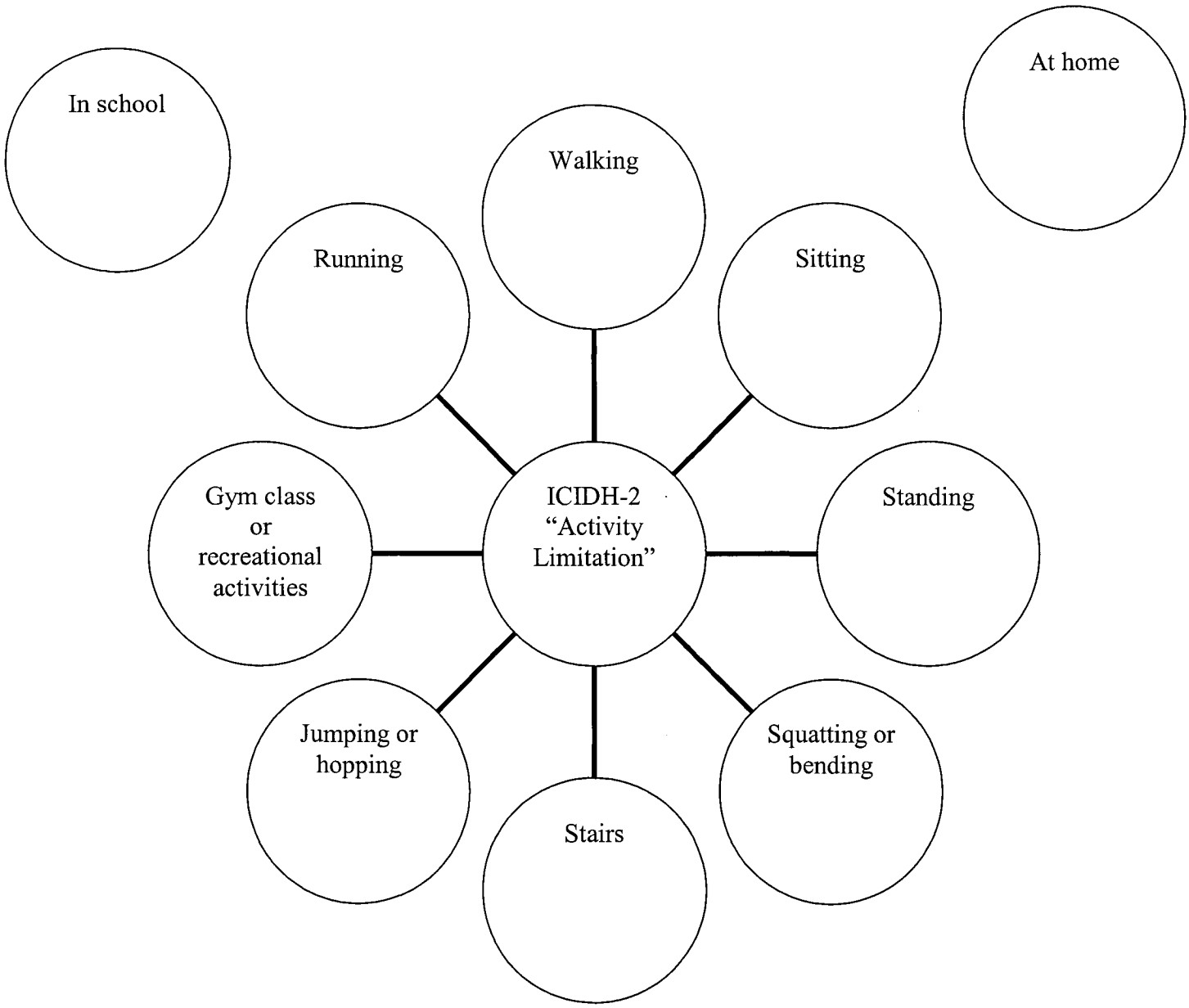
78. Stewart AL, Ware JE [eds]. Measuring Functioning and Well-Being: The Medical Outcomes Study Approach. Durham: Duke University Press, 1992.

79. Sapnas KG. Zeller RA. Minimizing sample size when using exploratory factor analysis for measurement. Journal of Nursing Measurement 2002; 10:135-54.

80. Froman RD. Elements to consider in planning the use of factor analysis. Southern Online Journal of Nursing Research 2001; 2:1-21.

81. Fayers PM. Machin D. Sample size: How many patients are necessary? British Journal of Cancer 1995; 72: 1-9.

Appendix 1 – Conceptual Framework used in the development of the Lower Extremity Activity Limitations Scale



Appendix 2 – Initial List of Self Report Items

Items	Eventual Fate of Item
Any of your usual work, housework, or school activities	Modified - #20 in final scale
Your usual hobbies, recreational or sporting activities	Modified - #21 in final scale
Getting into or out of the bath	Deleted – redundant
Walking between rooms	Deleted – redundant
Putting on your socks or shoes	#2 in final scale
Squatting	#3 in final scale
Lifting an object, like a bag of groceries off the floor	Modified - #4 in final scale
Performing light activities around your home	Deleted – redundancy/ face validity
Performing heavy activities around your home	Deleted – redundancy/ face validity
Getting into or out of a car	#6 in final scale
Walking 2 blocks	#8 in final scale
Walking for a mile	Modified - #9 in final scale
Going up or down 10 stairs (about 1 flight of stairs)	Modified - #10 in final scale
Standing for 1 hour	#12 in final scale
Sitting for 1 hour	#13 in final scale
Running on even ground	#14 on final scale
Running on uneven ground	#15 on final scale
Making sharp turns while running fast	#16 on final scale
Hopping	#18 on final scale
Rolling over in bed	Deleted – failed face validity
Used the toilet at home by myself	Deleted – failed face validity
Put on my pants by myself	Deleted – failed face validity
Put on my shoes and done them up by myself	Deleted- redundancy
Done my usual jobs or chores	Modified - #5 in final scale
Played the same sports and games that I usually enjoy by myself (Examples: gymnastics, jogging, swimming)	Supplemental #21
Walked without any support (No crutches or canes)	Deleted – redundant
Moved around inside my home without anyone to help me	Deleted – redundant
Walked in crowded areas	Deleted – redundant
Moved around outside without anyone to help me	Deleted – redundant
Walked up a gentle hill or slope by myself	Deleted – redundant
Walking on rough surfaces	Modified - #17 in final scale
Kept up with my friends outside	Deleted – redundant
Stood still for 10 minutes without resting	Deleted – redundant
Stretched to reach a high shelf (or to see over the person in front of me)	Deleted – failed face validity
Gotten through heavy doors by myself	Deleted – failed face validity
Walked up and down a flight of stairs	Deleted – redundant
Got into and out of an automobile	Deleted – redundant
Got into and out of a chair	Deleted – redundant
Sat on the floor	Deleted – redundant
Gotten in and out of my bed by myself	Deleted – redundant
Gotten down onto the floor or ground from standing and	Deleted – redundant

gotten back up again by myself	
Keeping up with my friends at school or at home	#19 in final scale
Participating in gym class	Supplemental #21
Walking around inside the house	#1 in final scale
Picking up things from the floor	Deleted – redundant
Climbing several flights of stairs	#11 in final scale
Jumping	Deleted – redundant
Getting into or out of a bus	#7 in final scale

Appendix 3 – Initial List of Functional Performance Test Items

Items	Eventual Fate of Item
Squatting items	
1. Double leg squat performed on a decline board	Deleted – redundant
2. Single leg squat on a flat surface	Deleted – redundant
3. Single leg squat on a decline board	Deleted – redundant
4. Deep knee bend	#1 in final scale
5. Repeated sit to stand	#2 in final scale
6. Double leg squat performed on a flat surface	Deleted – redundant
Jumping and hopping items	
7. Vertical jump	#3 in final scale
8. Standing hop	#5 in final scale
9. Standing two-foot jump	Deleted – redundant
10. One-legged jump	Deleted – redundant
11. 6-meter hop test	Deleted – redundant
12. 12-meter hop test	#6 in final scale
13. Cross over hop	Deleted – redundant
14. Triple hop for distance	Deleted – redundant
15. Stair hop	Deleted – redundant
16. Single leg hop on the decline board	Deleted – redundant
Stairs (Timed)	
17. Climbing 1 flight of stairs	Deleted – redundant
18. Step up	#4 in final scale
Walking and Running (Timed)	
19. 25 yard walk or run	Separated - #9 and #10 in final scale
20. Shuttle run	Deleted – redundant
21. Agility run	Deleted – redundant
22. Caricoa	Deleted – redundant
23. Semi-circular	Deleted – redundant
Balance (Timed)	
24. Static balance – eyes open/ closed	Separated - #7 and #8 in final scale

Appendix 4 – Information Letter

Information Letter

This letter is to inform you about a research study that is being performed in this department and to help you make an informed decision as to whether or not you want your child to participate in the study. Please read the letter thoroughly and feel free to ask any questions.

DEVELOPMENT OF A FUNCTIONAL SCALE FOR ADOLESCENTS WITH LOWER EXTREMITY ORTHOPEDIC CONDITIONS

Research Team
Sarah Milne (Physiotherapist and Masters' student, Department of Epidemiology and Community Medicine, University of Ottawa)
Marg Bisch, (Physiotherapist, Outpatient Physiotherapy Clinic, CHEO)
Dominique Surprenant, (Physiotherapist, Outpatient Physiotherapy Clinic, CHEO)
Brenda Elliott, (Physiotherapist, Outpatient Physiotherapy Clinic, CHEO)

Supervisors – Dr. Ian McDowell (Professor, University of Ottawa)
Gail Bowes (Professional Practice Leader, Physiotherapy Service, CHEO)

Why is the study being conducted?
When a child is seen in physiotherapy for an injury, before any treatment can begin, the physiotherapist must determine where there is a problem. In order to do this, the physiotherapist may try to measure things like your child's muscle flexibility, strength or amount of pain. These assessments are done routinely but we are currently looking for ways to improve them.

How many children will be in the study?
Up to 100 children from the outpatient physiotherapy clinic of the Children's Hospital of Eastern Ontario will be asked to participate.

You are eligible to participate if you are aged between 10 and 18 years old and have been referred for physiotherapy with any injury to the leg, knee or foot.

What will be involved if I choose to participate?
Your child will be recruited voluntarily from the outpatient physiotherapy clinic at CHEO. His/her physical function will be assessed using functional scale. The scale is made up of 2 parts: a questionnaire that is filled in by your child and a test that requires a physiotherapist to observe your child performing a series of common activities like

Children's Hospital of Eastern Ontario
Centre hospitalier pour enfants de l'est de l'Ontario

Lettre de renseignements

La présente lettre vise à vous informer sur une étude menée par le département et à vous aider à prendre une décision éclairée sur la participation de votre enfant à celle-ci. Veuillez lire attentivement ce qui suit et ne pas hésiter à poser des questions.

MISE AU POINT D'UNE ÉCHELLE FONCTIONNELLE POUR LES ADOLESCENTS ATTEINTS DE TROUBLES ORTHOPÉDIQUES DES MEMBRES INFÉRIEURS

Équipe de recherche
Sarah Milne (physiothérapeute et étudiante à la maîtrise, Département d'épidémiologie et de médecine sociale, Université d'Ottawa)
Marg Bisch (physiothérapeute, service des consultations externes en physiothérapie, CHEO)
Brenda Elliott (physiothérapeute, service des consultations externes en physiothérapie, CHEO)
Dominique Surprenant (physiothérapeute, service des consultations externes en physiothérapie, CHEO)

Directeurs de la recherche
Ian McDowell (professeur, Université d'Ottawa)
Gail Bowes (chef de pratique professionnel, service de physiothérapie, CHEO)

Quelle est la raison d'être de l'étude?
Lorsqu'un enfant consulte un physiothérapeute pour une blessure, ce dernier doit déterminer, avant d'amorcer le traitement, où se trouve le problème. Pour ce faire, il peut essayer de mesurer chez l'enfant des aspects comme sa souplesse musculaire, sa force ou l'ampleur de sa douleur. Ces évaluations sont pratique courante, mais nous sommes à la recherche de façons de les améliorer.

Combien d'enfants participeront à l'étude?
Jusqu'à 100 enfants du service des consultations externes en physiothérapie du CHEO seront pressentis pour participer à l'étude.

Peuvent participer à l'étude les jeunes âgés de 10 à 18 ans qui doivent faire de la physiothérapie à la suite d'une blessure à une jambe, à un genou ou à un pied.

En quoi consiste l'évaluation?
Votre enfant sera recruté sur une base volontaire au service des consultations externes en physiothérapie du CHEO. Sa fonction physique sera évaluée à l'aide d'une échelle fonctionnelle. Celle-ci comprend deux volets : un questionnaire que votre enfant devra

Appendix 5 – Consent Form

Children's Hospital of Eastern Ontario
Centre hospitalier pour enfants de l'est de l'Ontario

DEVELOPMENT OF A FUNCTIONAL SCALE FOR ADOLESCENTS WITH LOWER EXTREMITY ORTHOPEDIC CONDITIONS

We are studying whether our lower extremity functional scale is a good tool in detecting changes in adolescents with leg injuries. Although many scales have been developed for adults to measure physical function, a single scale has not been found and widely used in children.

Your child will be asked to complete the evaluation a maximum of 4 times. The evaluation will be administered during his/her initial assessment and during follow-up visits. The assessment will require your child to fill in 2 questionnaires and be evaluated by an experienced physiotherapist. The evaluation consists of a 21-item questionnaire related to the function of the legs and a physical test in which your child will be asked to perform various activities like walking, running and jumping. Your child will also be asked to complete a questionnaire on his/her quality of life. The entire evaluation will take approximately 20 minutes.

There will be no immediate benefit to your child for participating in the study. The results will however improve how we assess children such as yours. There are minimal risks involved if you are eligible to participate in this study. The evaluation is not harmful or painful. It does not involve any activities that are not routinely performed during a physiotherapy assessment. However, your child will be asked to answer several additional questions on some of the activities he/she does throughout the day and the assessment could produce the usual fatigue associated with activity.

All of the information obtained from this research will be used only to determine if the lower extremity functional scale is a good tool to use in children such as yours. I have been assured that personal records relating to this study will be kept strictly confidential except as required and permitted by law. Government regulators such as Health Canada, as well as representatives from the CHEO Research Board can have access to your child's personal information. I have been assured that my child will not be identified in any publication or presentation of this study. In addition, any personal information about my child that leaves the hospital will be coded so that my child cannot be identified by name.

Your participation in this study is entirely voluntary. You have the right to refuse participation or withdraw from the study at any time without reason or penalty. The regular treatment of your child will not be affected in any way. Your child will not have his/her care altered in any way other than agreeing to be evaluated using a standard form. No money is being offered for your child's involvement in this study.

Children's Hospital of Eastern Ontario
Centre hospitalier pour enfants de l'est de l'Ontario

DEVELOPMENT OF A FUNCTIONAL SCALE FOR ADOLESCENTS WITH LOWER EXTREMITY ORTHOPEDIC CONDITIONS

Adolescent Form

To be completed by adolescents aged 16 to 18 years

We are studying whether our lower extremity functional scale is a good tool in detecting changes in adolescents with leg injuries. Although many scales have been developed for adults to measure physical function, a single scale has not been found and widely used in children.

You will be asked to complete the evaluation a maximum of 4 times. The evaluation will be administered during the initial assessment and during follow-up visits. The assessment will require you to fill in 2 questionnaires and be evaluated by an experienced physiotherapist. The evaluation consists of a 21-item questionnaire related to the function of the legs and a physical test in which you will be asked to perform various activities like walking, running and jumping. You will also be asked to complete a questionnaire on my quality of life. The entire evaluation will take approximately 20 minutes.

There will be no immediate benefit for participating in the study. The results will however improve how we assess patients like you. There are minimal risks involved if you are eligible to participate in this study. The evaluation is not harmful or painful. It does not involve any activities that are not routinely performed during a physiotherapy assessment. However, you will be asked to answer several additional questions on some of the activities you do throughout the day and the assessment could produce the usual fatigue associated with activity.

All of the information obtained from this research will be used only to determine if the lower extremity functional scale is a good tool to use in patients like you. I have been assured that personal records relating to this study will be kept strictly confidential except as required and permitted by law. Government regulators such as Health Canada, as well as representatives from the CHEO Research Board can have access to your personal information. I have been assured that my I will not be identified in any publication or presentation of this study. In addition, any personal information about me that leaves the hospital will be coded so that I cannot be identified by name.

Your participation in this study is entirely voluntary. You have the right to refuse participation or withdraw from the study at any time without reason or penalty. Your regular treatment will not be affected in any way. You will not have your care altered in any way other than agreeing to be evaluated using a standard form. No money is being offered for your involvement in this study.

Children's Hospital of Eastern Ontario
Centre hospitalier pour enfants de l'est de l'Ontario

MISE AU POINT D'UNE ÉCHELLE FONCTIONNELLE POUR LES ADOLESCENTS ATTEINTS DE TROUBLES ORTHOPÉDIQUES DES MEMBRES INFÉRIEURS

Nous cherchons actuellement à déterminer, au moyen d'une étude, si notre échelle fonctionnelle des membres inférieurs est un outil adéquat pour détecter les changements chez les adolescents ayant des blessures aux jambes. Bien qu'un grand nombre d'échelles existent pour mesurer les fonctions physiques chez les adultes, aucune n'a été conçue et n'est largement utilisée pour les enfants.

L'évaluation se fera en un maximum de quatre séances, soit pendant la visite initiale puis au cours des visites de suivi. Votre enfant sera amené à remplir deux questionnaires et à rencontrer un physiothérapeute chevronné. L'évaluation consiste en un questionnaire qui compte 21 points ayant rapport à la fonction des jambes et en un test physique au cours duquel divers mouvements doivent être exécutés, comme marcher, courir et sauter. Votre enfant devra également remplir un questionnaire sur sa qualité de vie. En tout, l'évaluation prendra une vingtaine de minutes.

Votre enfant ne retirera aucun avantage immédiat de sa participation à l'étude. Toutefois, les conclusions de la recherche permettront d'améliorer la façon dont nous évaluons des enfants comme le vôtre. Les risques associés à l'étude sont minimes; l'évaluation n'est pas douloureuse et ne comporte aucune activité qui ne soit pas pratiquée couramment pendant une évaluation de physiothérapie. Cependant, votre enfant devra répondre à plusieurs questions supplémentaires sur certaines des activités qu'il accomplit pendant la journée, et l'évaluation pourrait entraîner la fatigue habituelle associée aux activités en question.

Toutes les données recueillies au cours de l'étude serviront exclusivement à déterminer si l'échelle fonctionnelle pour les membres inférieurs est un bon outil pour des enfants comme le vôtre. Les renseignements personnels recueillis pour cette étude seront gardés strictement confidentiels, sauf si la loi exige ou permet leur divulgation. Les organes de réglementation du gouvernement comme Santé Canada et les représentants du conseil d'administration de la recherche du Centre hospitalier pour enfants de l'est de l'Ontario (CHEO) pourront toutefois avoir accès à ces renseignements.

Vous êtes entièrement libre de participer à cette étude. Vous avez le droit de refuser d'y prendre part ou de vous retirer de celle-ci en tout temps sans avoir à justifier votre décision et sans subir de préjudices. Les soins habituels fournis à votre enfant ne seront modifiés d'aucune manière, à part sa participation à cette évaluation réalisée au moyen de formulaires standard. Aucune rétribution ne vous sera accordée pour avoir permis la participation de votre enfant à l'étude.
J'ai reçu l'assurance que mon enfant ne sera identifié dans aucune publication ni présentation liée à cette étude. En outre, tous les renseignements personnels le concernant

Children's Hospital of Eastern Ontario
Centre hospitalier pour enfants de l'est de l'Ontario

MISE AU POINT D'UNE ÉCHELLE FONCTIONNELLE POUR LES ADOLESCENTS ATTEINTS DE TROUBLES ORTHOPÉDIQUES DES MEMBRES INFÉRIEURS

Formulaire à l'intention des adolescents
À remplir par les jeunes de 16 à 18 ans

Nous cherchons actuellement à déterminer, au moyen d'une étude, si notre échelle fonctionnelle des membres inférieurs est un outil adéquat pour détecter les changements chez les adolescents ayant des blessures aux jambes. Bien qu'un grand nombre d'échelles existent pour mesurer les fonctions physiques chez les adultes, aucune n'a été conçue et n'est largement utilisée pour les enfants.

L'évaluation se fera en un maximum de quatre séances, soit pendant votre visite initiale puis au cours des visites de suivi. Vous remplirez deux questionnaires et rencontrerez un physiothérapeute chevronné. L'évaluation consiste en un questionnaire qui compte 21 points ayant rapport à la fonction des jambes et en un test physique au cours duquel divers mouvements doivent être exécutés, comme marcher, courir et sauter. Vous devrez également remplir un questionnaire sur votre qualité de vie. En tout, l'évaluation prendra une vingtaine de minutes.

Vous ne retirerez aucun avantage immédiat de votre participation à l'étude. Toutefois, les conclusions de la recherche permettront d'améliorer la façon dont nous évaluons des patients comme vous. Les risques associés à l'étude sont minimes; l'évaluation n'est pas douloureuse et ne comporte aucune activité qui ne soit pas pratiquée couramment pendant une évaluation de physiothérapie. Cependant, vous devrez répondre à plusieurs questions supplémentaires sur certaines des activités que vous accomplissez pendant la journée, et l'évaluation pourrait entraîner la fatigue habituelle associée aux activités en question.

Toutes les données recueillies au cours de l'étude serviront exclusivement à déterminer si l'échelle fonctionnelle pour les membres inférieurs est un bon outil pour des patients comme vous. Les renseignements personnels recueillis pour cette étude seront gardés strictement confidentiels, sauf si la loi exige ou permet leur divulgation. Les organes de réglementation du gouvernement comme Santé Canada et les représentants du conseil d'administration de la recherche du Centre hospitalier pour enfants de l'est de l'Ontario (CHEO) pourront toutefois avoir accès à ces renseignements.

Vous êtes entièrement libre de participer à cette étude. Vous avez le droit de refuser d'y prendre part ou de vous retirer de celle-ci en tout temps sans avoir à justifier votre décision et sans subir de préjudices. Les soins habituels qui vous sont fournis ne seront modifiés d'aucune manière, à part votre participation à cette évaluation réalisée au moyen de formulaires standard. Aucune rétribution ne vous sera accordée pour avoir participé à l'étude.

Appendix 6 – Self Report Booklet

Children's Hospital of Eastern Ontario
Centre hospitalier pour enfants de l'est de l'Ontario

Development and evaluation of a functional scale for use in adolescents with lower extremity orthopedic conditions

When people are sick or not feeling well it is sometimes difficult for them to do their regular activities. Please answer the following questions about your activities. If you have any questions while you are filling in the form, please ask your therapist. When you have finished filling out the form please hand it in to your therapist. Thank you for your time.

Name_____ Date_____

Today <u>do you</u> or <u>would you</u> have any difficulty with ..... (Circle your answer)

	I could not do this at all	With a big problem	With a moderate problem	With a little bit of a problem	I could do this with no problem
Walking around inside the house	0	1	2	3	4
Putting on my socks or shoes	0	1	2	3	4
Squatting	0	1	2	3	4
Lifting an object, like my backpack off the floor	0	1	2	3	4
Doing my usual jobs or chores at home	0	1	2	3	4
Getting into or out of a car	0	1	2	3	4
Getting into or out of a bus	0	1	2	3	4
Walking 2 blocks	0	1	2	3	4
Walking for 30 minutes	0	1	2	3	4
Going up or down a flight of stairs	0	1	2	3	4
Climbing several flights of stairs	0	1	2	3	4
Standing for 1 hour	0	1	2	3	4
Sitting for 1 hour	0	1	2	3	4
Running on even ground	0	1	2	3	4
Running on uneven ground	0	1	2	3	4
Making sharp turns while running fast	0	1	2	3	4
Walking on rough or slippery surfaces	0	1	2	3	4
Hopping	0	1	2	3	4
Keeping up with my friends at school or at home	0	1	2	3	4
Any of my usual school activities not including gym	0	1	2	3	4
My usual recreational or sporting activities including gym class	0	1	2	3	4

What activities or sports would you normally participate in at this time of the year?

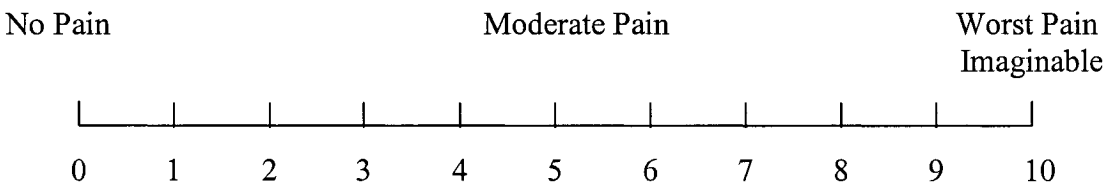
_____

_____

How many times a week would you normally do these activities? _____

Have you been able to continue these activities since you got hurt? _____

On average over the past two days, how much has your injury hurt?



KEEP GOING! YOU'RE ALMOST THERE.

During the past month, how often did you feel anxious, depressed, irritated, sad or downhearted and blue?

None of the time
A little of the time
Some of the time
Most of the time
All of the time

During the past month, if you needed someone to listen or to help you, was someone there for you?

Yes, as much as I wanted
Yes, quite a bit
Yes, some
Yes, a little
No, not at all

During the past month, how often did you talk to your problems, feelings or opinions with someone on your family?

All of the time
Most of the time
Some of the time
A little of the time
None of the time

Last week I think I could have	**I could not**	**With a big problem**	**With a moderate problem**	**With a little bit of a problem**	**With no problem**
Put toothpaste on my toothbrush then brushed my teeth by myself	0	1	2	3	4
Used the toilet at home by myself (includes getting on and off the toilet)	0	1	2	3	4
Washed my whole body by myself	0	1	2	3	4
Put my shirt on by myself	0	1	2	3	4
Put my pants on by myself	0	1	2	3	4
Fastened my clothes by myself	0	1	2	3	4
Put my shoes on and done them up by myself	0	1	2	3	4
Made a snack by myself	0	1	2	3	4
Done my usual jobs or chores (Examples: paper route, baby-sitting)	0	1	2	3	4
Taken care of my medical needs	0	1	2	3	4
Done my printing by myself	0	1	2	3	4
Played the same sports and games that I usually enjoy by myself (Examples: gymnastics, jogging)	0	1	2	3	4
Walked without any support (No crutches or canes)	0	1	2	3	4
Moved around inside my home without anyone to help me	0	1	2	3	4
Walked in crowded areas	0	1	2	3	4
Moved around outside without anyone to help me	0	1	2	3	4
Walked up a gentle hill or slope by myself	0	1	2	3	4
Walked on rough or slippery surfaces	0	1	2	3	4
Kept up with my friends outside when they were running	0	1	2	3	4
Carried a drink or food to the table by myself without spilling	0	1	2	3	4
Carried things in 2 hands by myself	0	1	2	3	4
Stood still for 10 minutes without resting	0	1	2	3	4
Stretched to reach a high shelf	0	1	2	3	4
Gotten through heavy doors by myself	0	1	2	3	4
Walked up and down a flight of stairs	0	1	2	3	4
Gotten in and out of an automobile by myself	0	1	2	3	4
Gotten in and out of a chair by myself	0	1	2	3	4
Sat on the floor	0	1	2	3	4
Gotten in and out of my bed by myself	0	1	2	3	4
Gotten down onto the floor or ground from standing and gotten back up again by myself	0	1	2	3	4

Appendix 7 – Therapist's Rating Form

Children's Hospital of Eastern Ontario
Centre hospitalier pour enfants de l'est de l'Ontario

Therapist Form

Development and evaluation of a functional scale for use in adolescents with lower extremity orthopedic conditions

Name _____ **Date** _____

Diagnosis _____

Visit

Initial Assessment ☐

Immediate post-op/ out of cast (Acute)	☐
1 month post-op/ out of cast, recent injury (Sub-acute)	☐
Chronic > 3 months (Chronic)	☐

< 72 hours post initial assessment ☐

Follow-up 3 weeks ☐

Follow-up 6 weeks ☐

Discharge ☐

Passive Range of Motion

	Right (degrees)	Left (degrees)	Comments
Hip			
Flexion			
Extension			
Internal Rotation (prone lying)			
External Rotation (prone lying)			
Knee			
Flexion			
Extension			
Ankle (long sitting)			
Dorsiflexion			
Plantar Flexion			

Strength

	Right (/5)	Left (/5)	Comments
Hip			
Flexors			
Adductors			
Abductors			
Extensors			
Internal Rotators			
External Rotators			
Knee			
Flexors			
Extensors			
90 degrees			
45 degrees			
0 degrees			
Ankle			
Dorsiflexors			
Plantar Flexors			
Inversion			
Eversion			

Flexibility

	Right (degrees)	Left (degrees)	Comments
Hip (Thomas Test position)			
Flexors			
Iliotibial Band			
Knee			
Quadriceps (prone lying)			
Hamstrings (SLR)			
Ankle (in standing)			
Gastrocnemius			
Soleus			

Additional Comments

_____

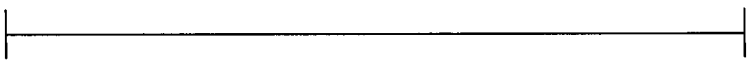
_____

_____

_____

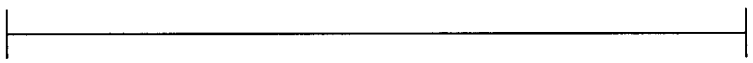
_____

From your experience with injuries of this nature, what impact would you expect this injury to have on the child's current activities of daily living?



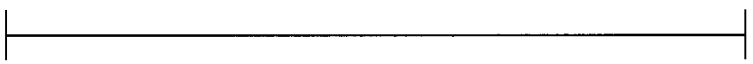
Severely disabling — Minimal Impact

From your interaction with this patient, how do you feel the patient is currently dealing with their injury?



In denial — Appropriate response — Exaggerated response

Where do you think this child is in terms of his/her recovery from his/her initial assessment? *(Not required at initial assessment)*



Much worse — Worse — No change — Improvement — Large Improvement

Additional Comments

_____

_____

_____

_____

_____

_____

_____

Appendix 8 – Lower Extremity Activity Limitations Scale – Functional Performance Test

Children's Hospital of Eastern Ontario
Centre hospitalier pour enfants de l'est de l'Ontario

Functional Performance Test Items

1. Double leg squat performed on a flat surface
   0 = unable
   1 = much difficulty (less than ½ knee ROM)
   2 = some difficulty (more than ½ knee ROM but not full OR with pain)
   3 = able

2. Repeated sit to stand with reach
   Subject is required to transition from sitting on a chair to standing and touching their dot, without using their hands. Count the number of repetitions performed in 30 seconds.

   Number of Repetitions _____

3. Two-foot static vertical jump

   Height of child _____ cm
   Height of jump _____ cm

4. One-legged standing hop

   Distance _____ cm

5. Step up
   Affected leg remains on a 10 inch step while the other leg steps up and down. Count the number of repetitions performed in 30 seconds.

   Number of Repetitions _____

6. 12-meter hop test

   Time to complete _____ seconds

7. One-legged static balance on affected leg
   Subject required to stand on affected leg with his/her **eyes open**. Time recorded. Maximum 30 seconds.

   Time _____ seconds

8. One-legged static balance on affected leg
   Subject required to stand on affected leg with his/her **eyes closed**. Time recorded. Maximum 30 seconds.

   Time _____ seconds

9. 22.5 meter walk

   Time to complete _____ seconds

10. 22.5 meter run

    Time to complete _____ seconds

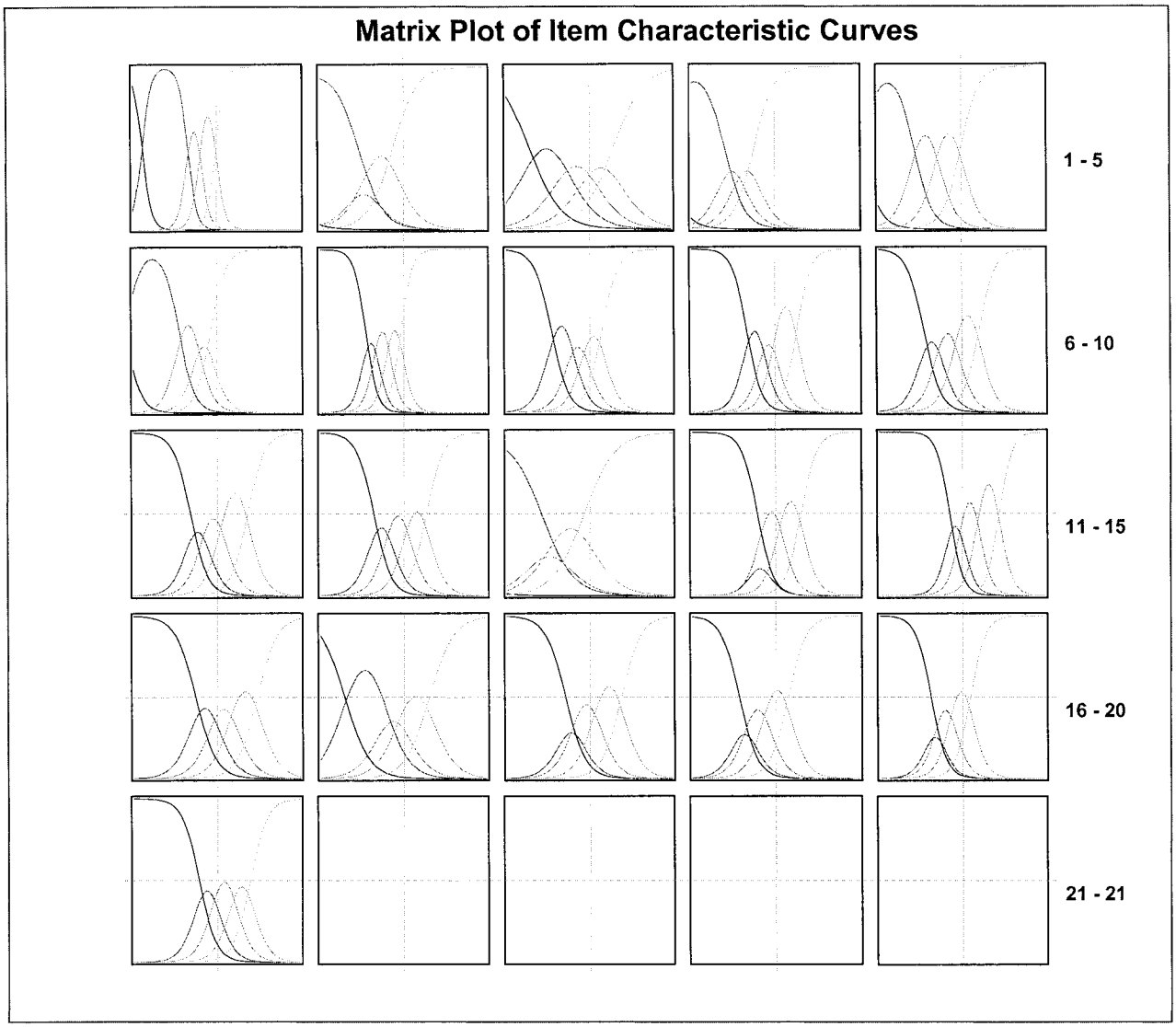
Appendix 9

Principal Components Analysis, Rotated Matrix, Oblimin Rotation (delta = 0)

	Factor				
	1	2	3	4	5
1. Walking around inside the house	.299	-.159	<0.10	-.613	<0.10
2. Putting on my socks or shoes	<0.10	.121	<0.10	**-.888**	-.224
3. Squatting	.246	-.268	-.243	<0.10	.206
4. Lifting an object, like my backpack off the floor	<0.10	.113	-.112	**-.515**	.191
5. Doing my usual jobs or chores at home	.102	-.212	<0.10	**-.556**	.116
6. Getting into or out of a car	<0.10	-.106	<0.10	**-.859**	<0.10
7. Getting into or out of a bus	<0.10	-.111	<0.10	**-.819**	<0.10
8. Walking 2 blocks	**.627**	-.186	<0.10	-.191	<0.10
9. Walking for 30 minutes	**.586**	-.136	-.235	-.171	<0.10
10. Going up or down a flight of stairs	**.589**	<0.10	-.203	-.225	.104
11. Climbing several flights of stairs	**.612**	<0.10	-.186	-.205	<0.10
12. Standing for 1 hour	**.544**	-.104	-.277	-.138	<0.10
13. Sitting for 1 hour	.153	<0.10	<0.10	-.392	**.426**
14. Running on even ground	<0.10	<0.10	**-.920**	<0.10	<0.10
15. Running on uneven ground	.242	<0.10	**-.873**	<0.10	<0.10
16. Making sharp turns while running fast	<0.10	<0.10	**-.889**	<0.10	<0.10
17. Walking on rough or slippery surfaces	<0.10	<0.10	**-.485**	-.275	.174
18. Hopping	<0.10	<0.10	**-.788**	-.180	<0.10
19. Keeping up with my friends at school or at home	**.461**	-.173	-.169	-.261	-.134
20. Any of my usual school activities not including gym	.281	<0.10	-.133	**-.484**	<0.10
21. My usual recreatinal or sporting activities including gym class	.179	-.128	**-.637**	-.107	<0.10
1. Deep knee bend	-.321	-.342	-.383	-.183	.284
2. Repeated sit to stand	.170	**-.720**	<0.10	<0.10	<0.10
3. Vertical jump	.139	**-.638**	<0.10	-.141	-.164
4. Distance of hop	<0.10	**-.401**	**-.437**	-.137	**-.491**
5. Step up	.239	**-.813**	<0.10	<0.10	<0.10
6. 12 meter hop	.295	.250	**.743**	<0.10	.363
7. Balance	-.176	**-.734**	<0.10	<0.10	<0.10
8. Proprioception	<0.10	**-.672**	<0.10	<0.10	-.128
9. 25 yard walk	<0.10	**.779**	<0.10	<0.10	-.182
10. 25 yard run	.181	**.574**	.313	.114	<0.10

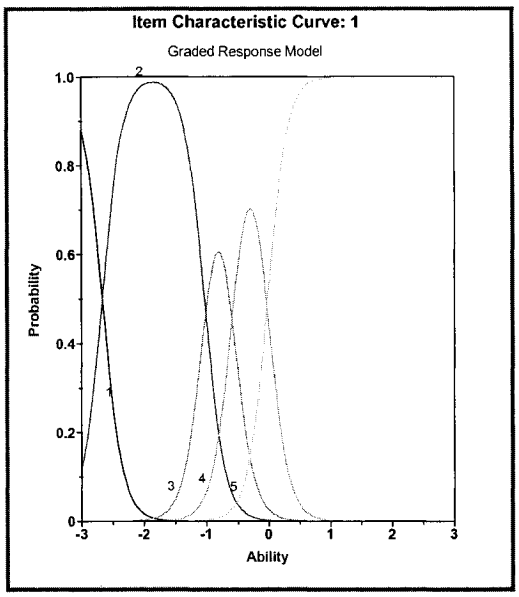
Appendix 10

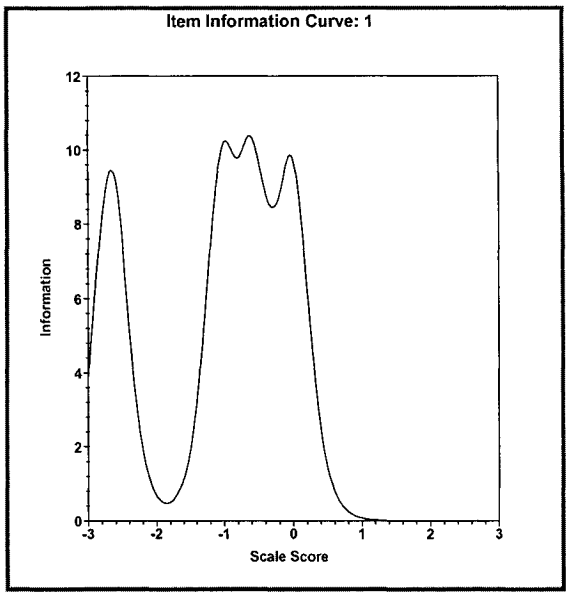
Item Response Theory, Item Characteristic Curves
LEALS- SR, 21- Items



Category legends

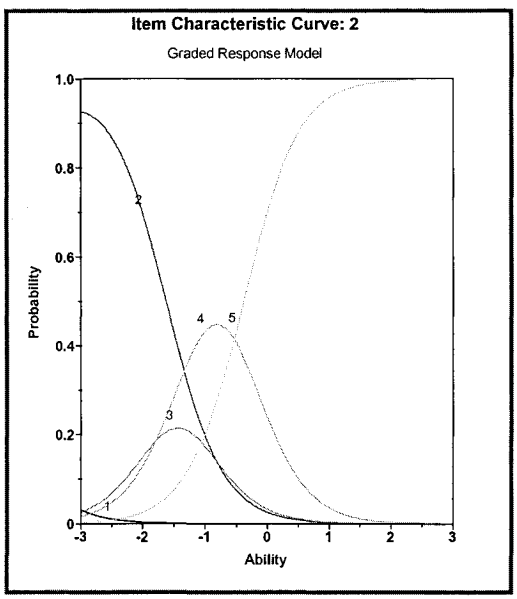
Solid Lines: 1= Black; 2 = Blue; 3 = Magenta; 4 = Green; 5 = Cyan

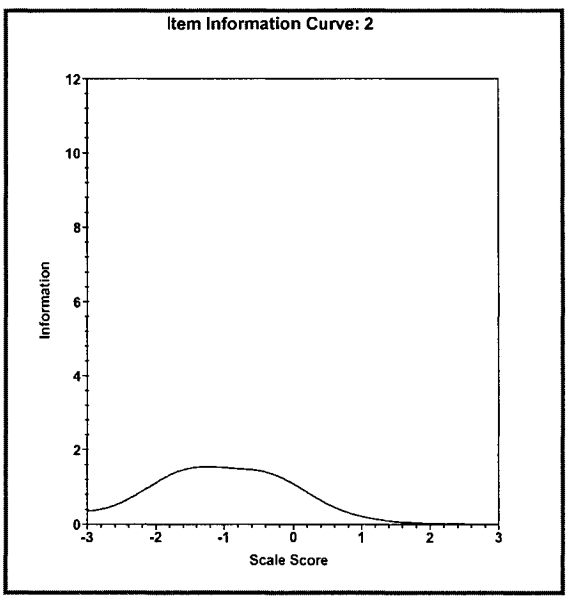




Category legends **Item: 1**

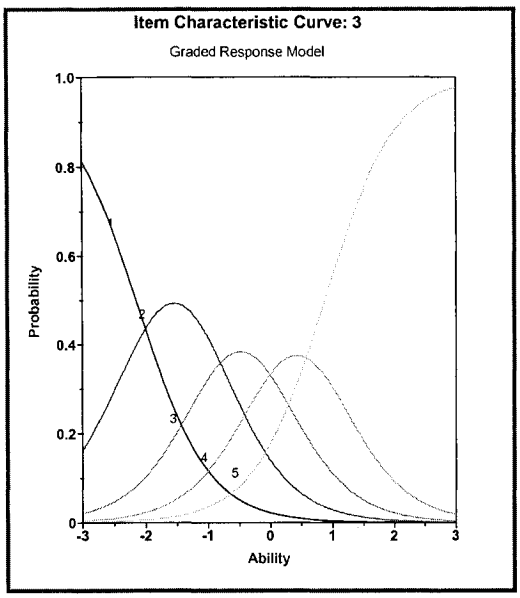
Solid Lines: 1= Black 2= Blue 3= Magenta 4= Green 5= Cyan

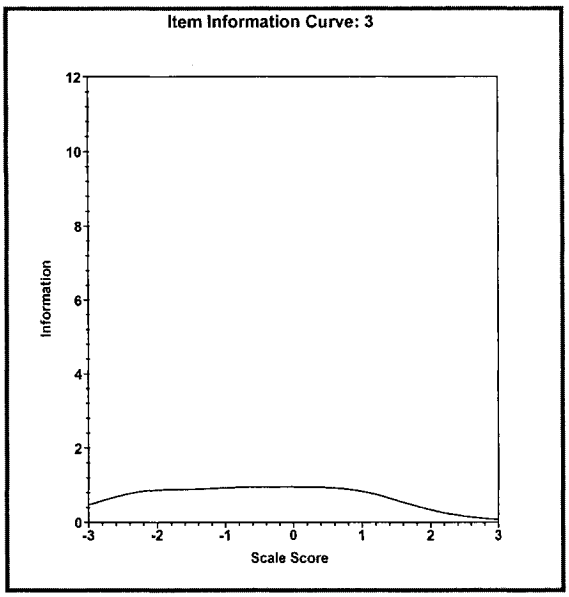




Category legends **Item: 2**

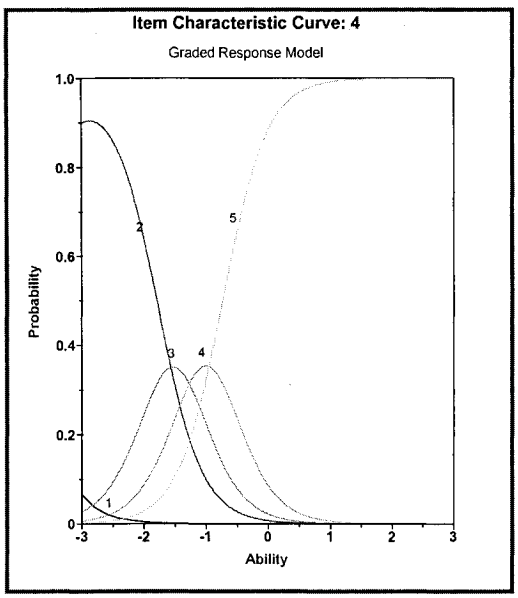
Solid Lines: 1= Black 2= Blue 3= Magenta 4= Green 5= Cyan

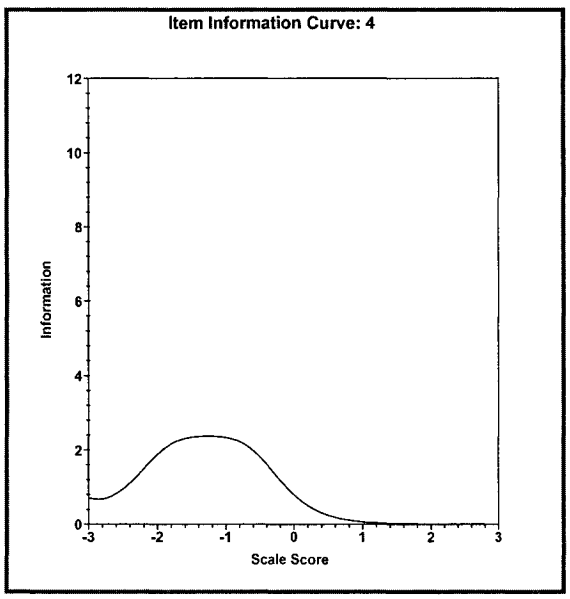




Category legends **Item: 3**

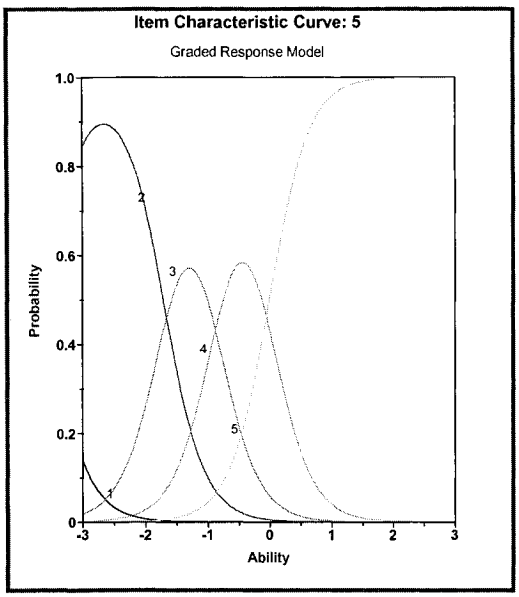
Solid Lines: 1= Black 2= Blue 3= Magenta 4= Green 5= Cyan

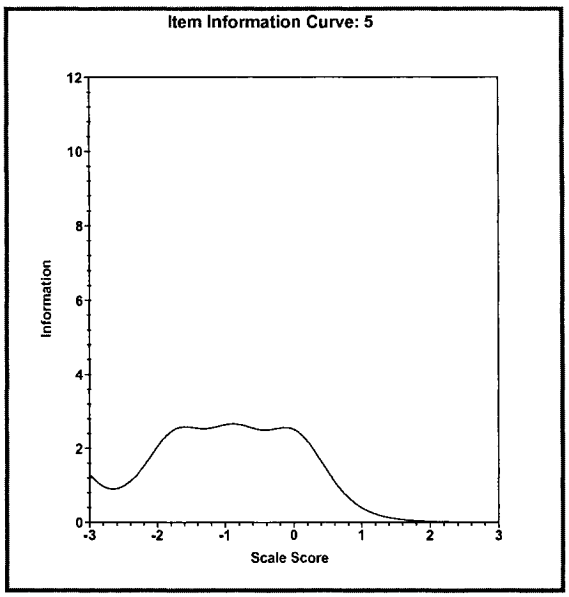




Category legends **Item: 4**

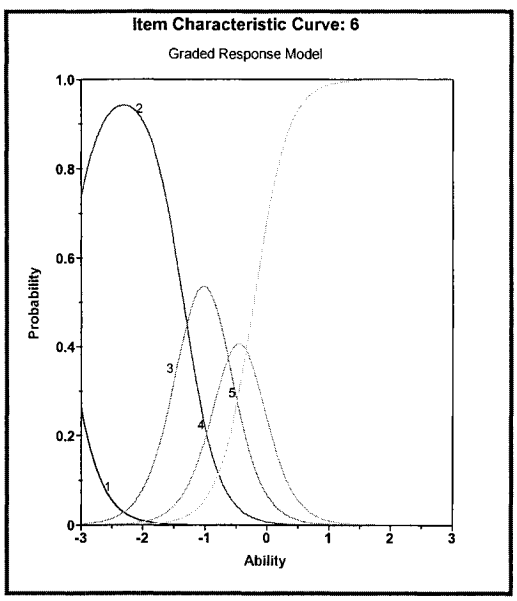
Solid Lines: 1= Black 2= Blue 3= Magenta 4= Green 5= Cyan

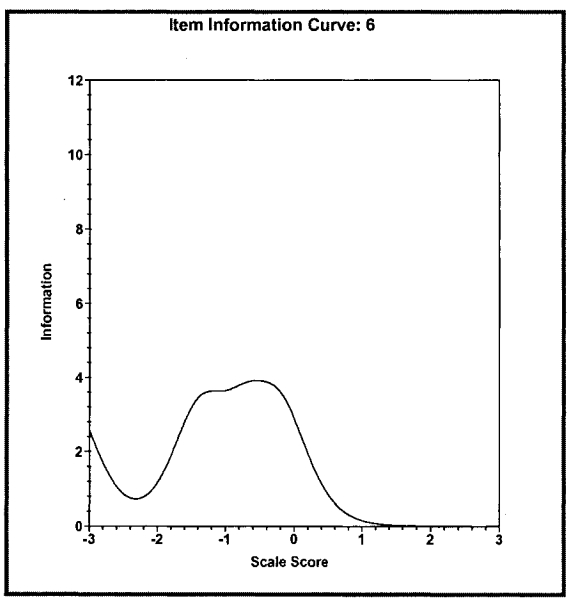




Category legends **Item: 5**

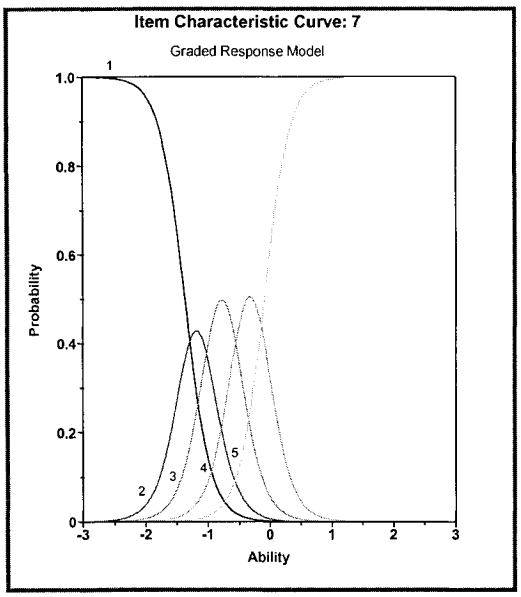
Solid Lines: 1= Black 2= Blue 3= Magenta 4= Green 5= Cyan

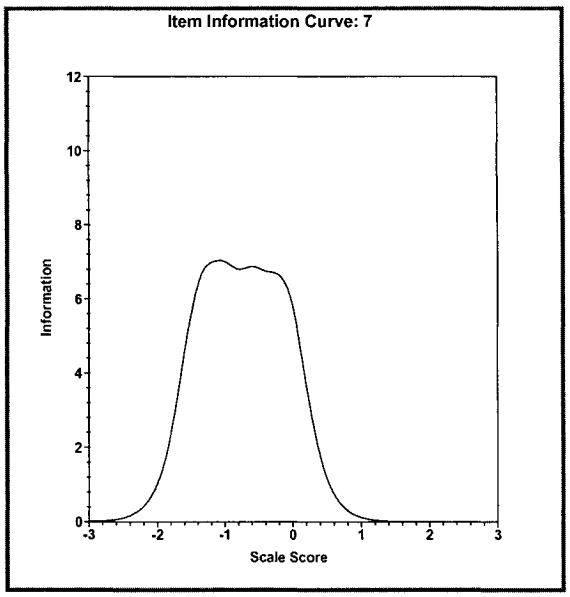




Category legends **Item: 6**

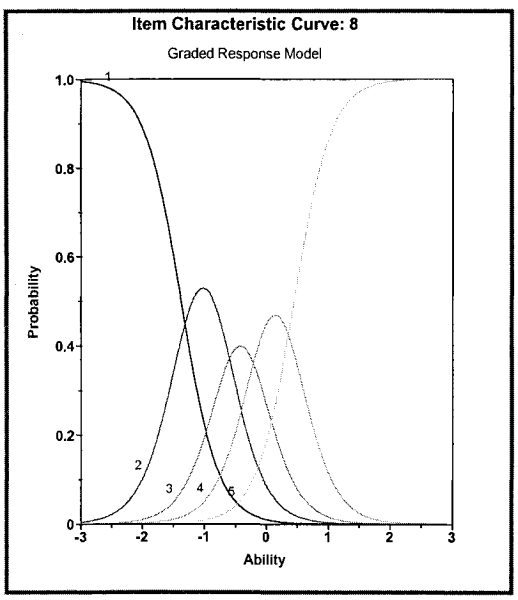
Solid Lines: 1= Black 2= Blue 3= Magenta 4= Green 5= Cyan

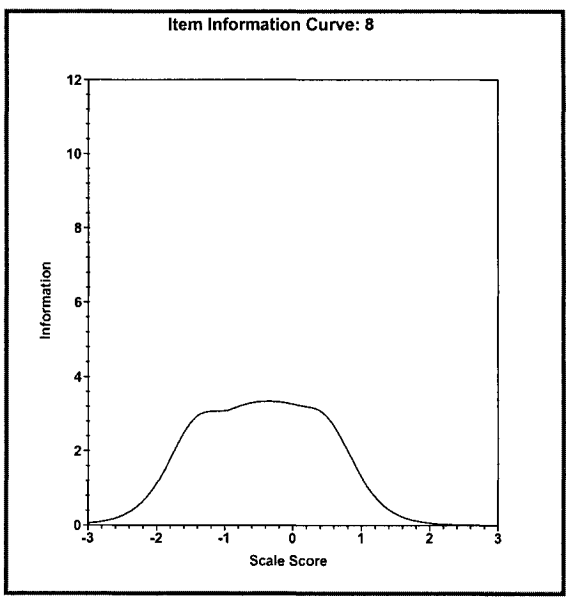




Category legends **Item: 7**

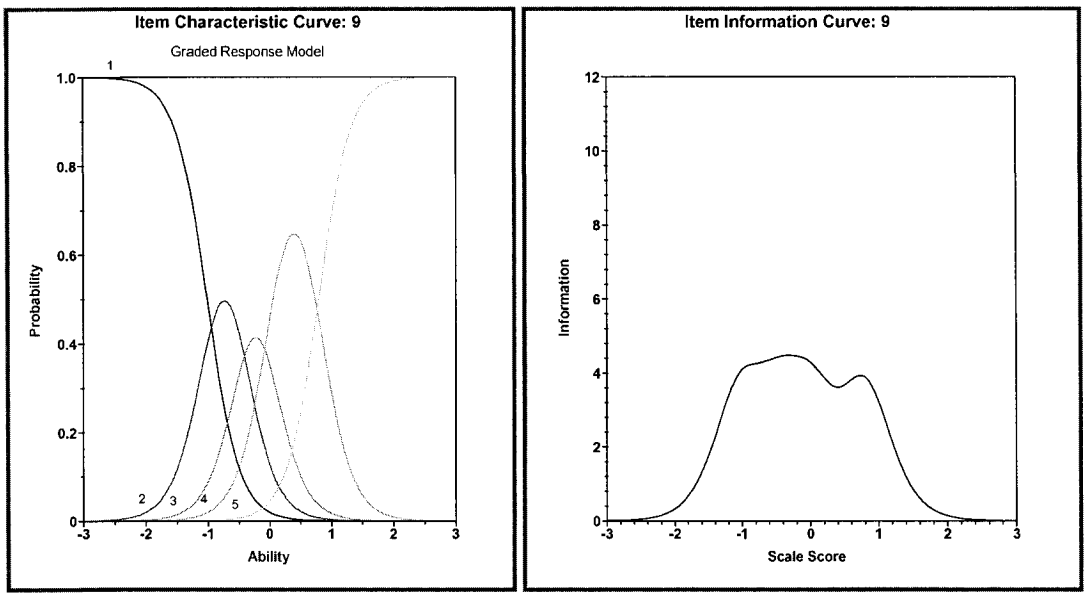
Solid Lines: 1= Black 2= Blue 3= Magenta 4= Green 5= Cyan





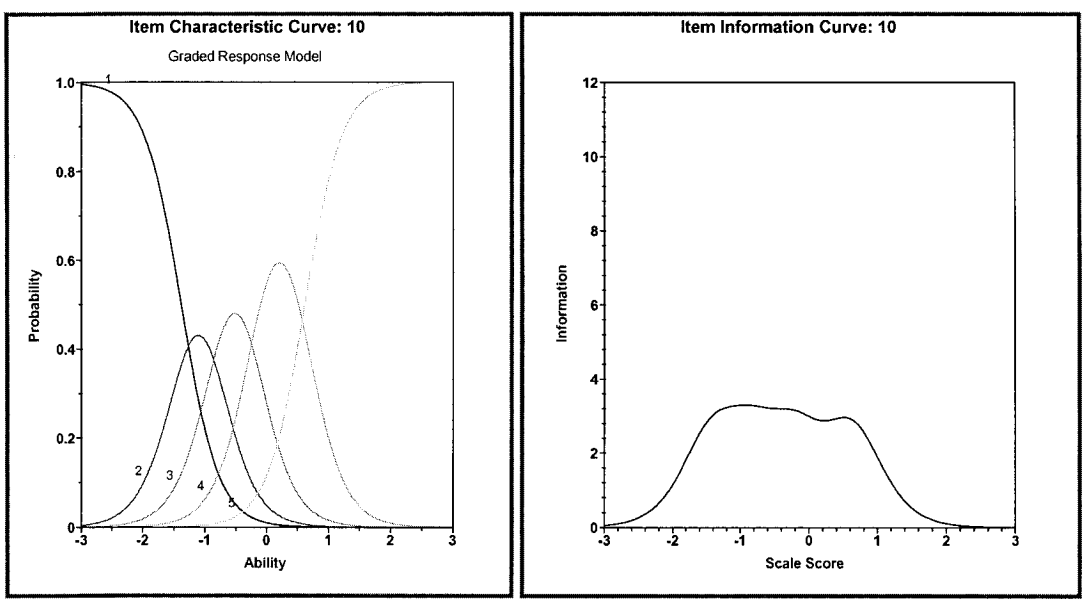
Category legends **Item: 8**

Solid Lines: 1= Black 2= Blue 3= Magenta 4= Green 5= Cyan



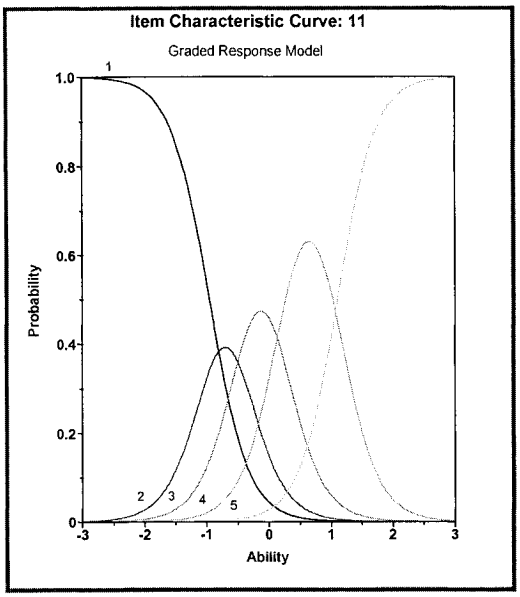
Category legends **Item: 9**

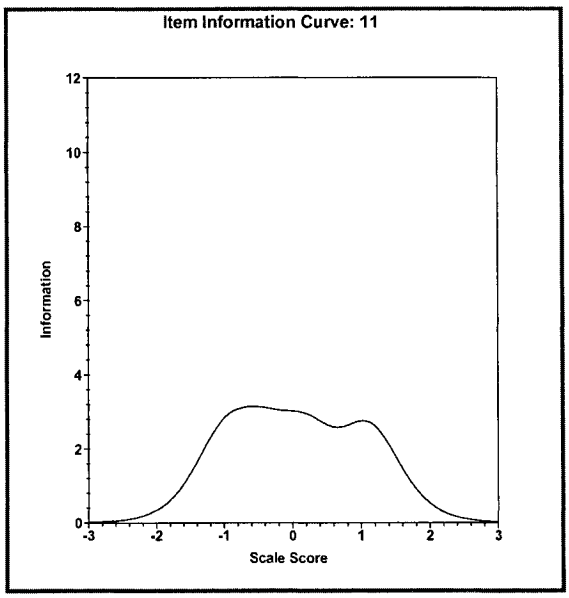
Solid Lines: 1= Black 2= Blue 3= Magenta 4= Green 5= Cyan



Category legends **Item: 10**

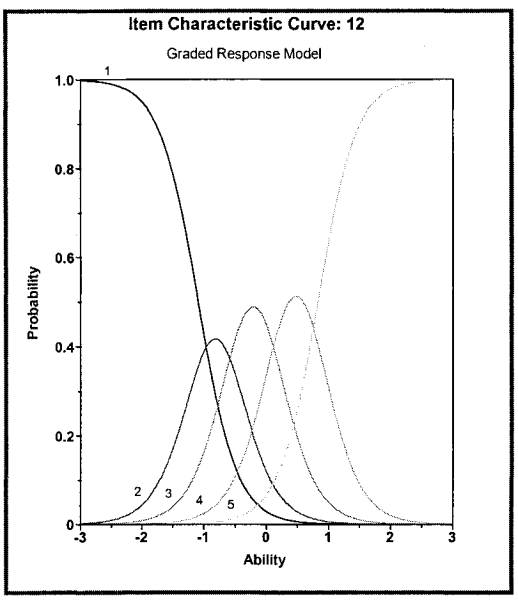
Solid Lines: 1= Black 2= Blue 3= Magenta 4= Green 5= Cyan

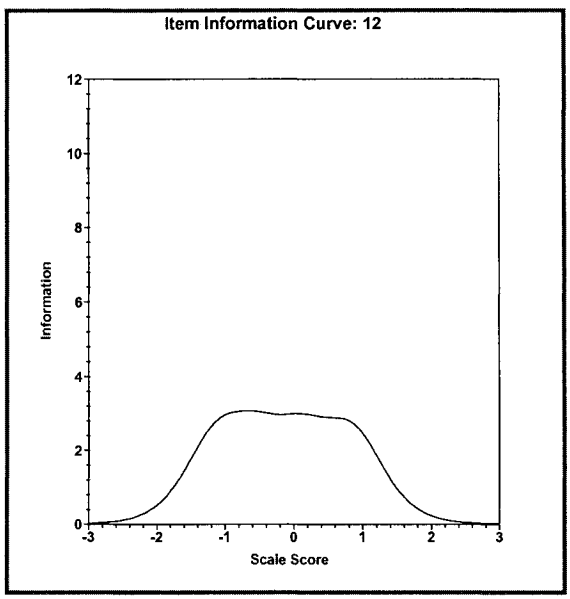




Category legends **Item: 11**

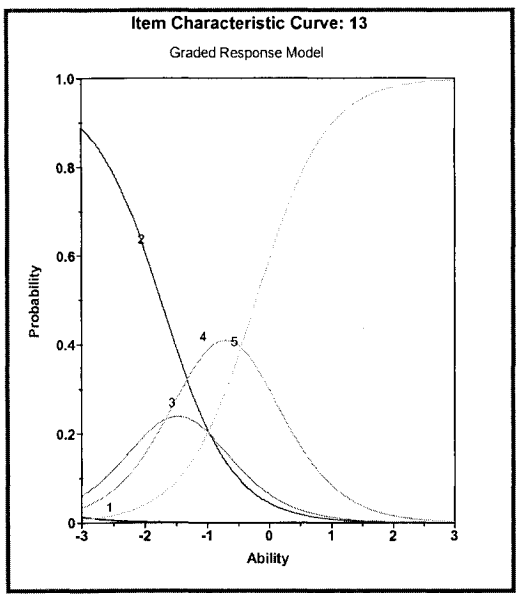
Solid Lines: 1= Black 2= Blue 3= Magenta 4= Green 5= Cyan

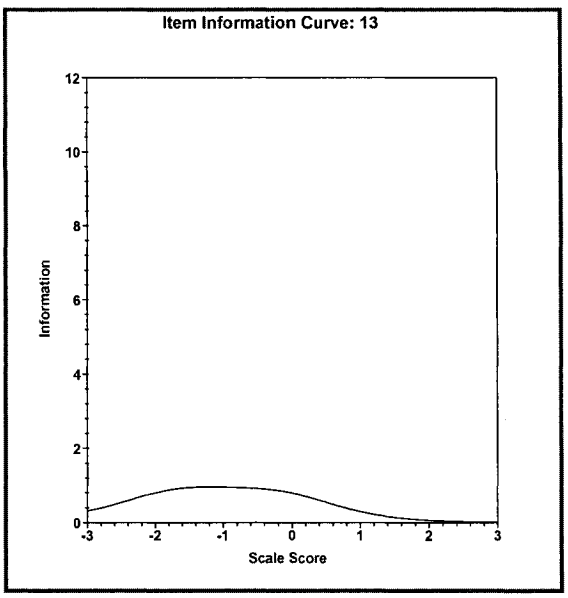




Category legends **Item: 12**

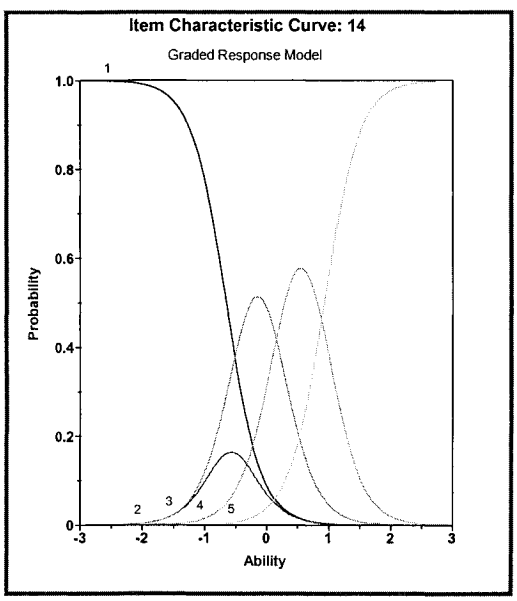
Solid Lines: 1= Black 2= Blue 3= Magenta 4= Green 5= Cyan

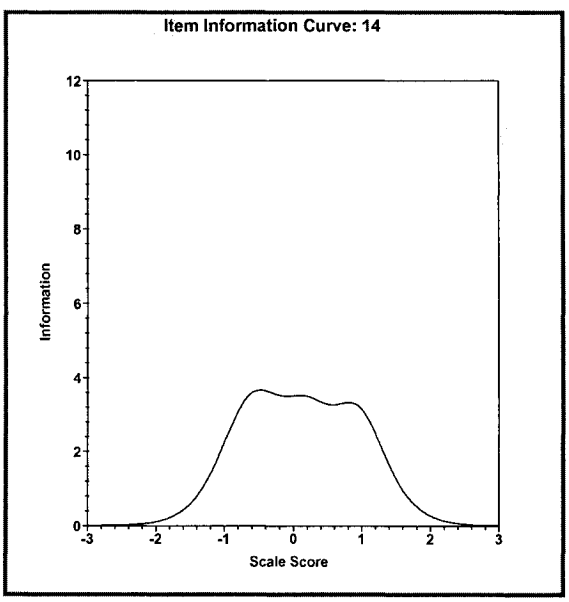




Category legends **Item: 13**

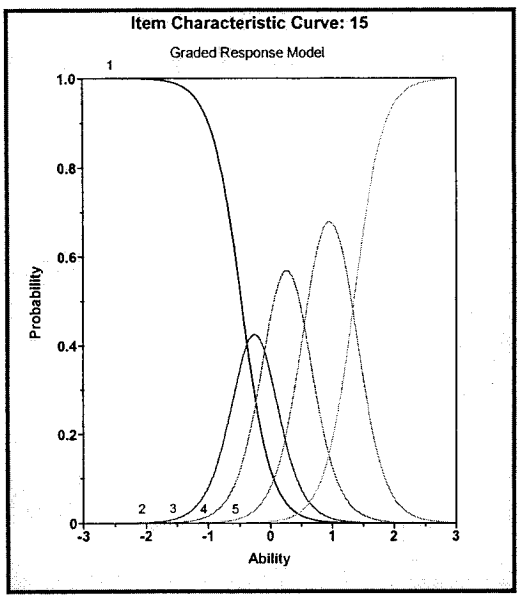
Solid Lines: 1= Black 2= Blue 3= Magenta 4= Green 5= Cyan

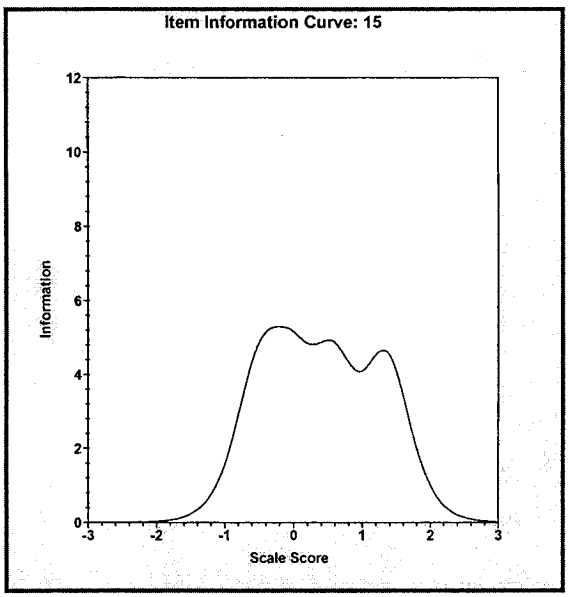




Category legends **Item: 14**

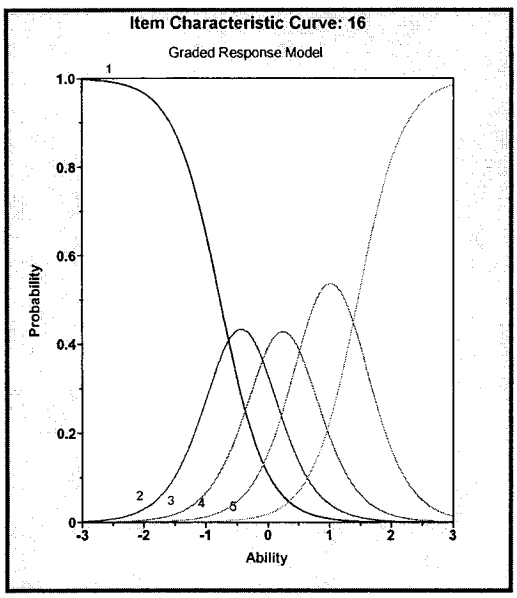
Solid Lines: 1= Black 2= Blue 3= Magenta 4= Green 5= Cyan

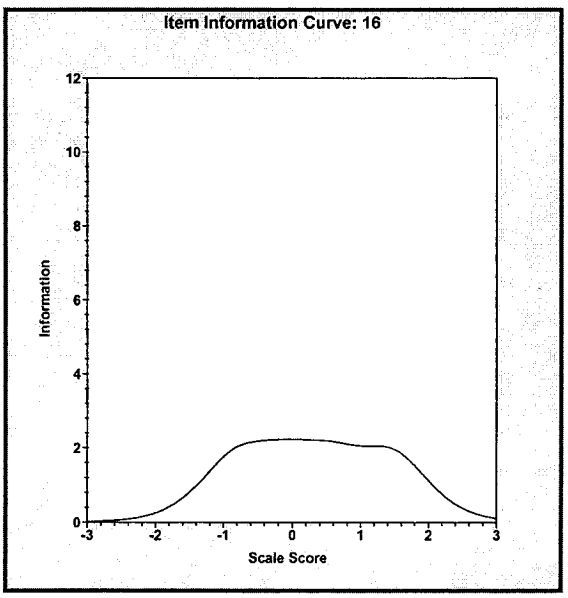




Category legends **Item: 15**

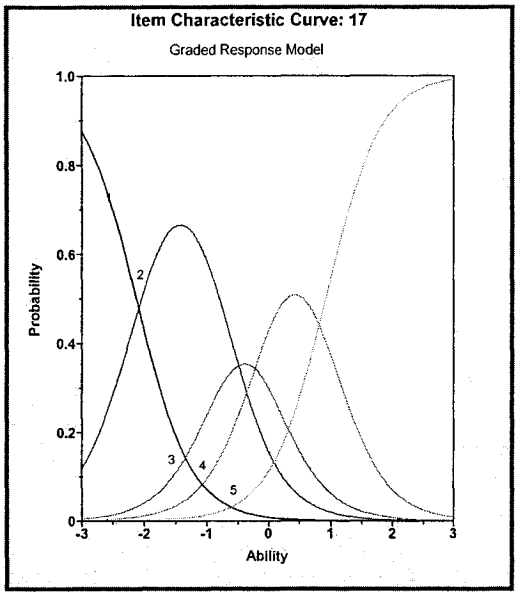
Solid Lines: 1= Black 2= Blue 3= Magenta 4= Green 5= Cyan

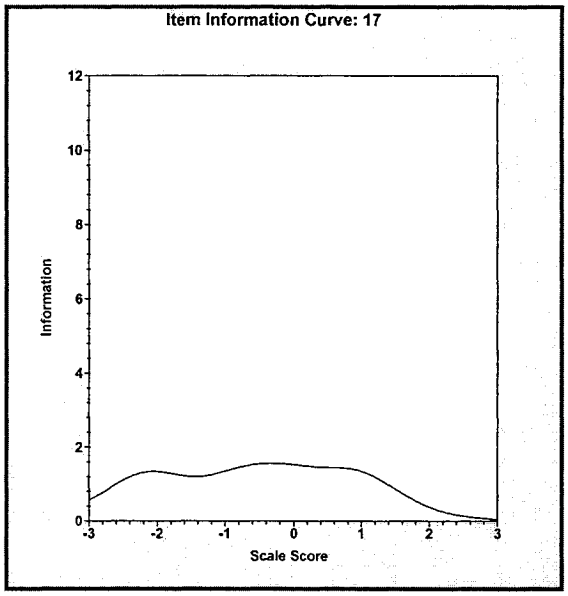




Category legends **Item: 16**

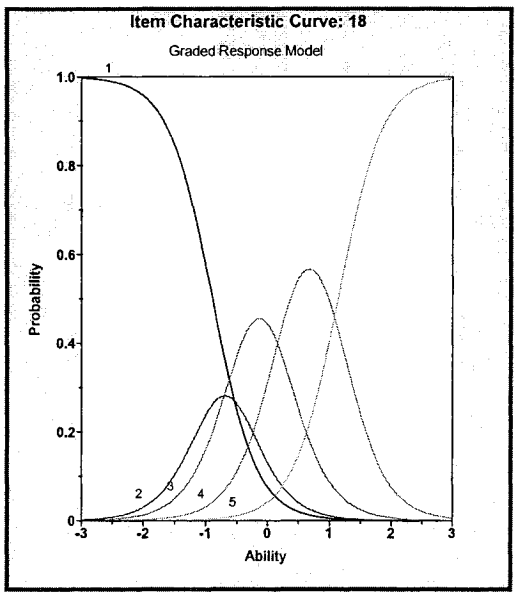
Solid Lines: 1= Black 2= Blue 3= Magenta 4= Green 5= Cyan

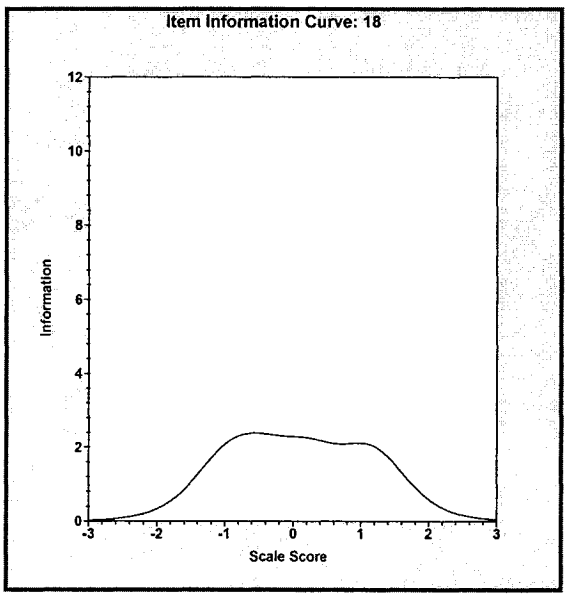




Category legends **Item: 17**

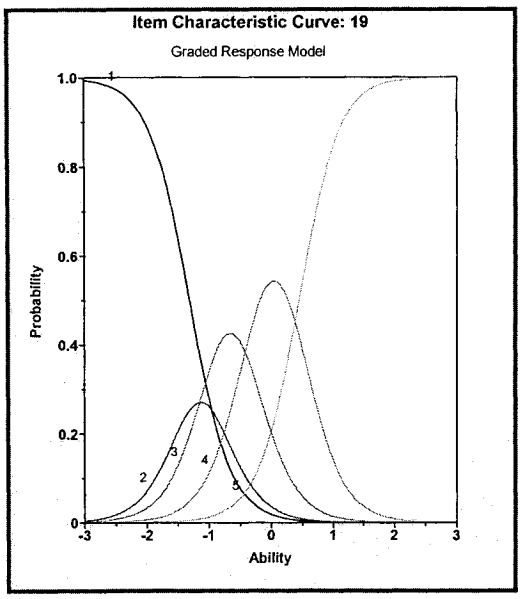
Solid Lines: 1= Black 2= Blue 3= Magenta 4= Green 5= Cyan

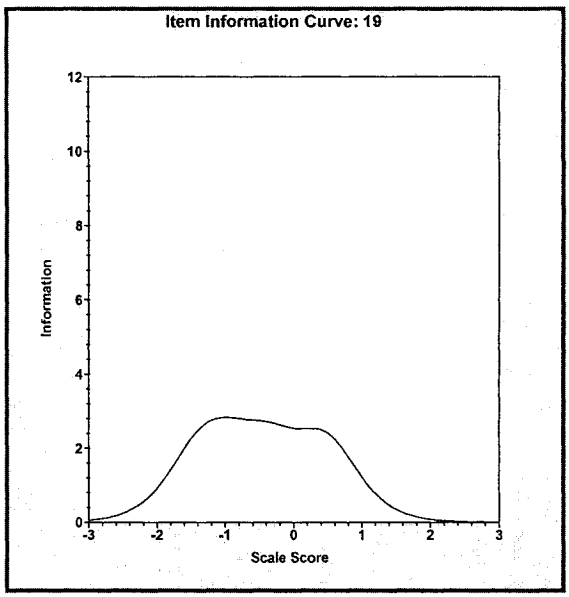




Category legends **Item: 18**

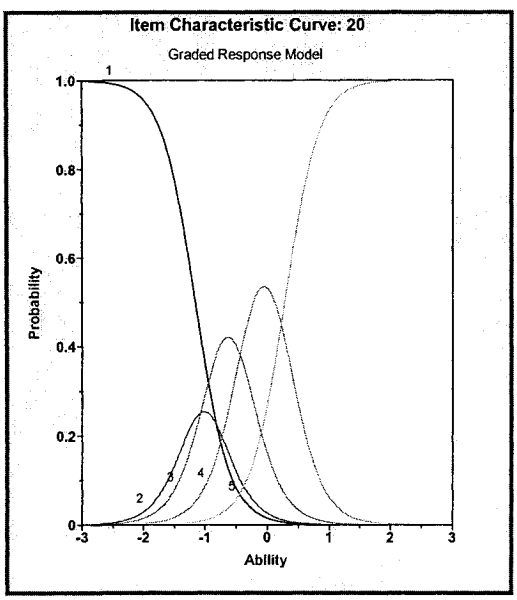
Solid Lines: 1= Black 2= Blue 3= Magenta 4= Green 5= Cyan

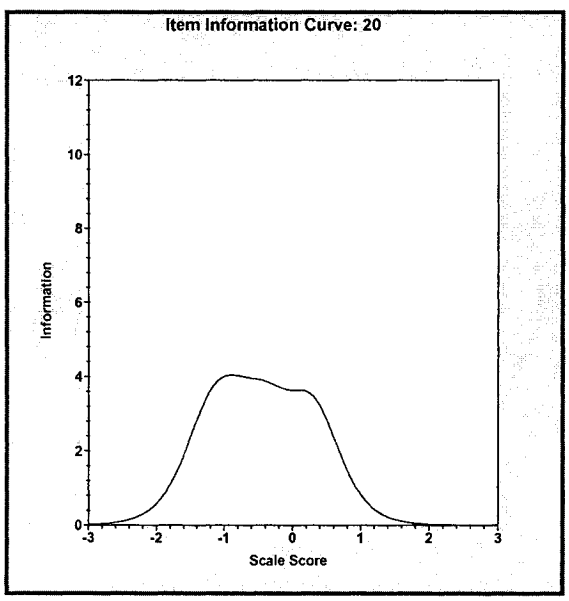




Category legends **Item: 19**

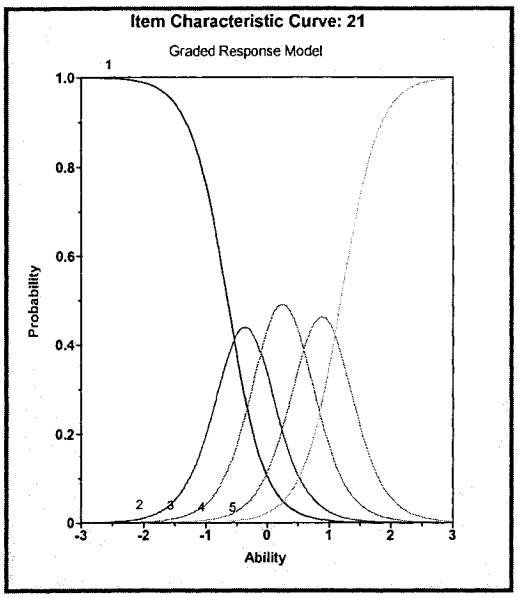
Solid Lines: 1= Black 2= Blue 3= Magenta 4= Green 5= Cyan

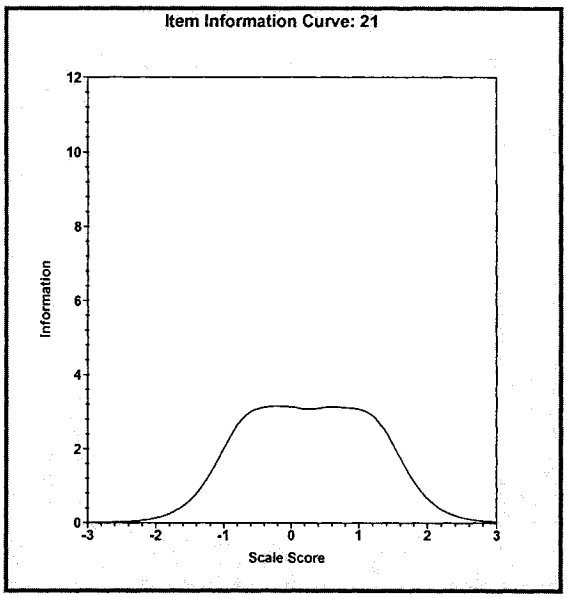




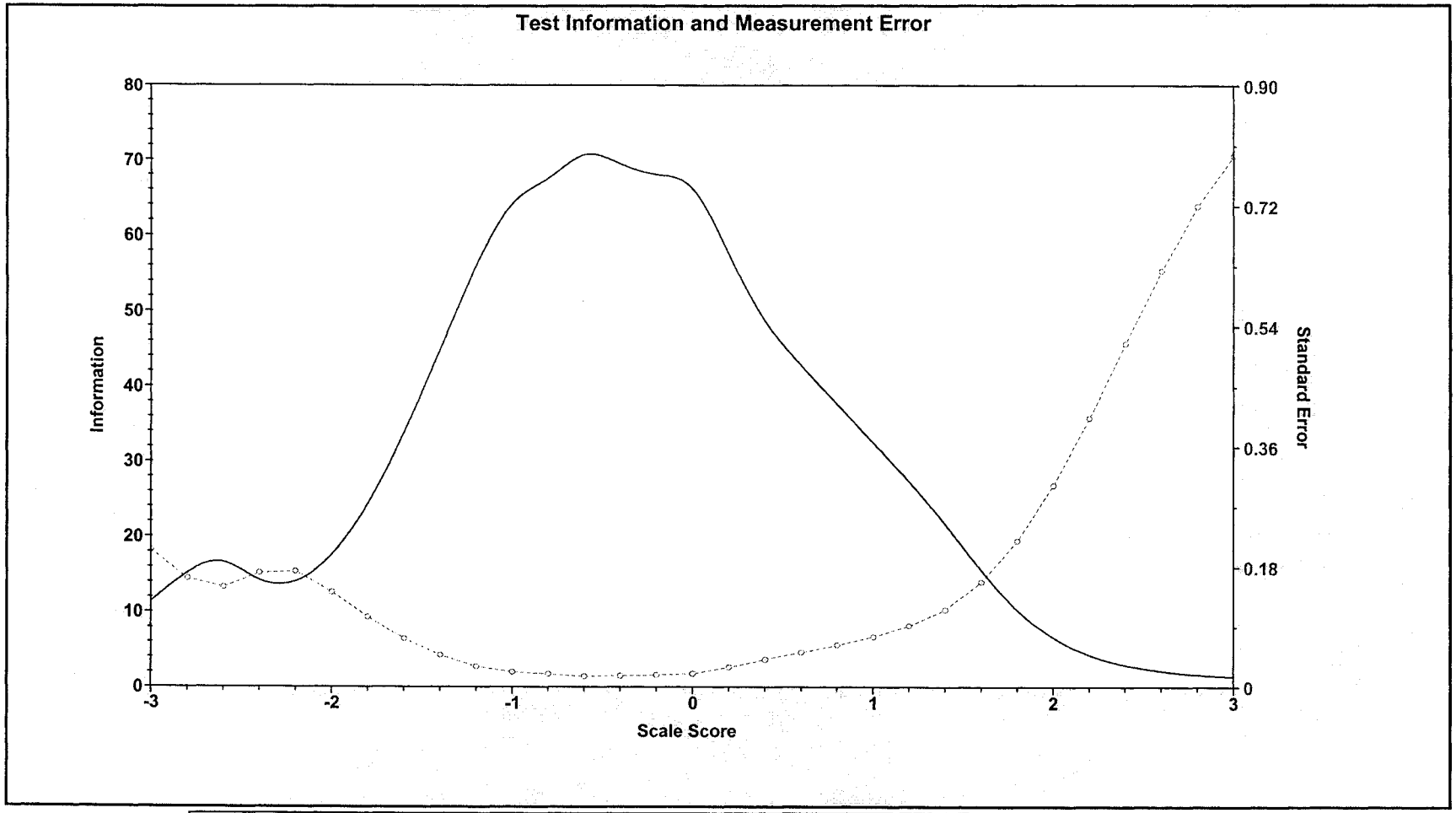
Category legends **Item: 20**

Solid Lines: 1= Black 2= Blue 3= Magenta 4= Green 5= Cyan





Category legends	Item: 21
Solid Lines: 1= Black 2= Blue 3= Magenta 4= Green 5= Cyan	



Test information curve: solid line **Standard error curve: dotted line**

The total test information for a specific scale score is read from the left vertical axis.

The standard error for a specific scale score is read from the right vertical axis.

Appendix 11 – Correlation Matrix (n = 120)

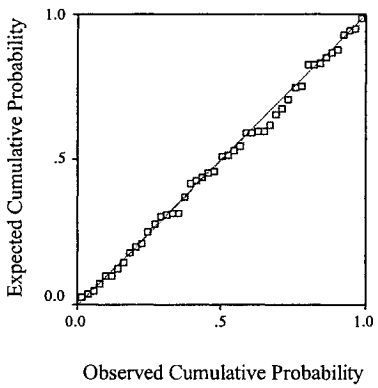
Variable	Self-report items	Performance items
Pain	-0.704**	-0.211*
Quality of life $^{\Delta}$	-0.223**	-0.228*
Range of motion	0.414**	0.595**
Strength	0.463**	0.662**
Impact VAS	0.432**	0.578**
Response VAS	-0.441**	-0.114

Pearson's correlation coefficient (2-tailed), ** correlation is significant at the 0.01 level (2-tailed), * correlation is significant at the 0.05 level, $^{\Delta}$ summary score of three questions combined

Normal P-P Plot of Regression

Standardized Residual

LEDS - SR



Normal P-P Plot of Regression

Standardized Residual

LEDS- Functional performance test