

IMPROVING PATIENT REPORTED MEASURES FOR KNEE FUNCTION

by

James L. Farnsworth II

A Dissertation Submitted in Partial Fulfillment
of the Requirements for the Degree of
Doctor of Philosophy in Human Performance

Middle Tennessee State University

December 2017

Dissertation Committee:

Dr. Minsoo Kang, Chair

Dr. Helen Binkley

Dr. Todd Evans

ACKNOWLEDGEMENTS

I would like to thank my wonderful wife Sarah for all of the sacrifices she has made to help me throughout this long process. Without her support, I would not have been able to accomplish my goal of finishing my degree. I would also like to give special thanks to Dr. Brian Ragan, who has been one of the greatest influences in my life. Although he is no longer with us, his mentorship has greatly impacted me. I will be forever grateful for the lessons he shared with me, which helped to shape me into the educator and researcher I am today. In addition, I would like to thank my advisor Dr. Minsoo Kang who has gave me this incredible opportunity to obtain my doctoral degree. His guidance has been a great asset to me as I worked on my degree and as I continue to advance my career as a professional in academia. Lastly, I would like to thank the members of my committee Dr. Todd Evans, and Dr. Helen Binkley for all of the assistance they have provided me throughout my dissertation process.

ABSTRACT

Knee injuries account for nearly 15% of sports-related injuries. Patient-reported outcomes (PRO) such as the KOOS and the IKDC-SKF are used to evaluate changes in knee-related function. Ceiling effects limit the clinical utility of PRO when used to monitor function. Therefore, the purpose of this project was to evaluate and improve PRO for athletes.

In the first study measurement properties of three PRO (e.g., KOOS, IKDC-SKF, and Marx Activity Rating Scale) were examined using the Rasch partial-credit model. Among the 65 items examined only 30 yielded acceptable model-data fit. Furthermore, measurement precision decreased as function increased. To improve functional assessment a new PRO measure was developed by combining items from the first study with 12 new items.

In the second study the psychometric properties of the new PRO measure were evaluated. Thirty-two student-athletes were recruited to pilot test the PRKOAT. Feedback regarding readability and relevance was used to modify the instrument. The revised instrument was calibrated using 203 student-athletes. Rasch partial-credit model analysis indicated that majority of items ($n = 27$) had acceptable model-data fit. Poor fitting items were removed.

Item difficulty ranged from -4.74 - 1.89 logits. Person ability ranged from -3.24 - 2.29. Injured athletes (mean score 39.25 ± 14.00), scored significantly higher (lower knee function) than non-injured athletes (11.93 ± 10.78) ($t_{188} = 12.89$; $p < .01$) providing evidence of known-group difference validity. Among

participants sampled only 3.9% obtained the best score (i.e., floor effect), while none of the participants obtained the worst score (i.e., ceiling effect).

A major advantage of this study was the use of advanced measurement theory. Compared to commonly used PRO measures the PRKOAT has fewer ceiling effects in athletic populations. Clinicians should consider using the PRKOAT to evaluate PRO throughout the rehabilitation process following knee-related injuries.

TABLE OF CONTENTS

LIST OF TABLES	x
LIST OF FIGURES	xi
CHAPTER I: INTRODUCTION	1
Sports Knee Injury Epidemiology	1
Consequences of Knee Injuries	2
Patient-Reported Outcomes	2
Patient-Reported Outcomes in Clinical Practice	3
Types of Patient-Reported Outcomes	3
Injury Evaluation Models	4
Knee-Related Patient-Reported Outcomes Instruments	5
Item Response Theory	6
Needs Statement	7
CHAPTER II: LITERATURE REVIEW	8
Sports Knee Injury Epidemiology	8
Consequences of Knee Injuries	10
Extended Lost Time from Competition	11
Consequences of Extended Time Lost Injuries	13
Injury Evaluation Models	14
Disablement Models	14

Functional Models.....	17
Historical Review of Knee-Related Injury Outcomes Measures	21
Larson Scoring Scale.....	21
Marshall System	22
Lysholm-Tegner Rating System.....	22
Lukianov Anterior Cruciate Ligament Evaluation Format.....	25
Cincinnati Knee Rating Scale	25
International Knee Documentation Committee	27
Patient-Reported Outcomes.....	30
Advantages of Patient-Reported Outcomes.....	31
Types of Patient-Reported Outcomes Instruments	33
Patient-Reported Outcomes Instruments for the Assessment of Knee-Related Function	35
Knee Injury and Osteoarthritis Outcomes Score.....	36
Marx Activity Rating Scale	42
International Knee Documentation Committee Subjective Knee Form	45
Limitations of Patient-Reported Outcomes Instruments	49
Content Validity.....	49
Population-Specific Validity.....	51
Ceiling Effects	53
Prevalence of Ceiling Effects	55
Performance-Based Outcomes: A Measure of Higher-Level Function? ..	56

Improving Patient-Reported Outcomes.....	56
Test Construction & Item Development.....	57
Item Response Theory.....	57
Fundamentals of Item Response Theory	57
Item Response Theory Model Characteristics and Assumptions.....	66
Item and Test Information Functions	73
Relative Efficiency	77
Advantages of Item Response Theory over Classical Test Theory	77
Historical Origins of Item Response Theory and Rasch Measurement Models	79
Philosophical Differences between Item Response Theory and Rasch ...	81
Rasch Rating Scale Model.....	82
Rasch Partial Credit Model	84
Summary	85
Specific Aims	86
CHAPTER III: PSYCHOMETRIC EVALUATION OF PATIENT-REPORTED KNEE OUTCOMES FOR ASSESSMENT OF KNEE-RELATED FUNCTION ...	88
Methods	92
Participants and Procedures.....	92
Instruments	94
Data Analysis.....	95
Results	97

Discussion	104
Limitations.....	108
Conclusion	108
References	109
CHAPTER IV: DEVELOPMENT AND VALIDATION OF THE PATIENT	
REPORTED KNEE OUTCOMES ASSESSMENT TOOL (PRKOAT)	114
Methods	117
Phase 1: Pilot Test.....	117
Phase 2: Rasch Calibration	130
Results	132
Discussion	140
Limitations.....	144
Conclusion	145
References	146
CHAPTER V: OVERALL CONCLUSION.....	151
REFERENCES	155
APPENDICES	182
APPENDIX A: IRB APPROVAL LETTER CHAPTER III	183
APPENDIX B: RECRUITMENT SCRIPT CHAPTER III.....	185
APPENDIX C: COMBINED KNEE FUNCTION ASSESSMENT TOOL	186

APPENDIX D: IRB APPROVAL LETTER CHAPTER IV	204
APPENDIX E: CONSENT FORM CHAPTER IV	205
APPENDIX F: RECRUITMENT SCRIPT CHAPTER IV.....	207
APPENDIX G: RECRUITMENT FLYER CHAPTER IV	208
APPENDIX H: PRKOAT (Pilot Version)	209
APPENDIX I: PRKOAT (Revised)	218

LIST OF TABLES

Table 1. Prevalence of Knee-Related Injuries for Each Sport	10
Table 2. Prevalence of Extended Lost Time Knee-Related Injuries	12
Table 3. Summary of the Types of Patient Reported Outcomes Instruments	34
Table 4. International Knee Documentation Committee Subjective Knee Form Questions Domains and Content.....	46
Table 5. Descriptive Statistics.....	93
Table 6. Patient Reported Outcomes – Knee Scoring Summary.....	98
Table 7. Rasch Partial-Credit Model Item Statistics for Final Model.....	100
Table 8. Item Statistics – Pilot Test	120
Table 9. PRKOAT Pilot Test Modifications	123
Table 10. Participant Demographic Information – Phase 2	131
Table 11. Misfitting Items.....	133
Table 12. Item Statistics Rasch Partial-Credit Model for Reduced Item Model.	135

LIST OF FIGURES

Figure 1. Nagi's Disablement Model.....	15
Figure 2. International Classification of Impairments, Disabilities, and Handicaps.....	17
Figure 3. World Health Organization International Classification of Function, Disability, and Health.....	20
Figure 4. International Knee Documentation Committee "Ideal" Evaluation and Documentation Guidelines	28
Figure 5. Theoretical Knee-Related Function Distribution	54
Figure 6. Item Response Theory Item Characteristic Curve	60
Figure 7. Item Characteristic Curve (1-parameter model)	62
Figure 8. Item Characteristic Curve (2-parameter model)	64
Figure 9. Item Characteristic Curve (3-parameter model)	66
Figure 10. An item characteristic curve and ability distribution for two groups of examinees	72
Figure 11. Item Information Function for Four Items.....	75
Figure 12. Test Information Function for Four Items.....	76
Figure 13. Illustration of Item-Parameters and Person-Ability Estimates Placed on the Same Scale	79
Figure 14. Rasch Rating Scale Model Category Probability Curve.....	84
Figure 15. Theoretical Knee-Related Function Distribution	90

Figure 16. Test Information Function and Distribution of Person-Ability Estimates.....	103
Figure 17. Conditional Standard Error of Combined Patient-Reported Knee Outcomes Measures	104
Figure 18. Score Distribution for the Patient Reported Knee Outcomes Assessment Tool (PRKOAT).....	138
Figure 19. Conditional Standard Error of Measurement for the PRKOAT	139

CHAPTER I

INTRODUCTION

Sports Knee Injury Epidemiology

The National Collegiate Athletic Association (NCAA) estimates that 178,925 injuries occurred between 1988-2004 across 14 intercollegiate sports (Agel, Dick, Nelson, Marshall, & Dompier, 2007; Agel, Dompier, Dick, & Marshall, 2007; Agel, Evans, Dick, Putukian, & Marshall, 2007; Agel, Olson, et al., 2007; Agel, Palmieri-Smith, Dick, Wojtys, & Marshall, 2007; Agel, Ransone, Dick, Oppliger, & Marshall, 2007; Dick, Ferrara, et al., 2007; Dick, Hertel, Agel, Grossman, & Marshall, 2007; Dick, Lincoln, et al., 2007; Dick, Putukian, Agel, Evans, & Marshall, 2007; Dick, Romani, Agel, Case, & Marshall, 2007; Dick, Sauers, Agel, & Keuter, 2007; S. W. Marshall, Covassin, Dick, Nassar, & Agel, 2007; S. W. Marshall, Hamstra-Wright, Dick, Grove, & Agel, 2007). While the majority of these injuries resolved quickly with minimal complications, approximately one fourth of these injuries resulted in extended lost time from competition (i.e., ten or more days) (Agel, Dick, et al., 2007; Agel, Dompier, et al., 2007; Agel, Evans, et al., 2007; Agel, Olson, et al., 2007; Agel, Palmieri-Smith, et al., 2007; Agel, Ransone, et al., 2007; Dick, Ferrara, et al., 2007; Dick, Hertel, et al., 2007; Dick, Lincoln, et al., 2007; Dick, Putukian, et al., 2007; Dick, Romani, et al., 2007; Dick, Sauers, et al., 2007; S. W. Marshall, Covassin, et al., 2007; S. W. Marshall, Hamstra-Wright, et al., 2007). Among injuries resulting in extended time loss, 25% were knee-related. Because of the importance of the knee in the

kinetic chain these injuries can have significant long-term complications if not treated properly.

Consequences of Knee Injuries

Adults with a history of knee injury often report lower health-related quality of life following injury (Lam, Thomas, Valier, McLeod, & Bay, 2017). A cross-sectional study of adults indicated that those with prior anterior cruciate ligament injuries were more likely to develop osteoarthritis and significant knee pain, despite surgical interventions (L. S. Lohmander, Englund, Dahl, & Roos, 2007). Because of the potential for long-term complications following knee injuries, it is vital that appropriate documentation is used to ensure patients are receiving the most optimal care.

Patient-Reported Outcomes

Patient-reported outcomes (PRO) are measurements related to a patient's health status from the perspective of the patient. There are many advantages to using PRO compared to traditional clinician based measures (Espallargues, Valderas, & Alonso, 2000; Greenhalgh & Meadows, 1999; Lohr, 1992; McHorney, 1999). Firstly, these instruments can help to facilitate the detection of physical and psychological problems that might be overlooked during treatment. In addition, these instruments can be used to measure changes in health status throughout the recovery process. Another benefit of PRO is that they may facilitate patient-clinician communication promoting shared decision making.

Patient-Reported Outcomes in Clinical Practice

Although there are still some barriers associated with implementation of PRO in clinical practice (Deyo & Patrick, 1989), many professional health organizations such as the National Athletic Trainers Association (NATA) have been encouraging clinicians to include PRO instruments as part of their clinical practice to monitor the recovery of patients (Lam, 2016; McLeod et al., 2008; Snyder et al., 2008). Currently, the NATA education competencies includes four competencies specific to patient outcomes assessment (NATA Executive Committee for Education, 2011). In addition, Medicare services has implemented new policies, beginning in 2017, which will require PROs for reimbursement of medical services. For these reasons, it is important for clinicians, specifically athletic trainers (ATs), to be familiar with these measures and understand the advantages and disadvantages associated with their use, to ensure that they are implemented correctly into clinical practice.

Types of Patient-Reported Outcomes

There are more than 50 different PRO instruments available for use in clinical practice. Selection of PRO is dependent upon the needs of the patient and the clinician. There are two primary types of PRO instruments, which include generic and specific instruments. Generic PRO instruments include items that assess broad aspects of health, while, specific PRO instruments include items related to a specific health condition, body region, dimension of health, or population.

The targeted nature of specific PRO instruments allows for increased diagnostic accuracy and responsiveness compared with generic PRO instruments for a specific health condition (Bombardier et al., 1995; Kantz, Harris, Levitsky, Ware Jr, & Davies, 1992). Because of their targeted nature, many specific PRO instruments have limited use beyond their specific health condition, body region or population. For example, the Knee Injury and Osteoarthritis Outcomes Score (KOOS), an instrument designed to assess knee function, does not include any items for the upper extremity. One advantage of generic PRO instruments, however, is because they are not specific to any one health condition, comparisons can be made across health conditions and populations.

Injury Evaluation Models

A number of injury evaluation models, such as the Nagi disablement model (S. Nagi, 1965), and the International Classification of Function, Disability, and Health (ICF) model (World Health Organization, 2001) have been established to serve as frameworks facilitating the development of new PRO instruments. The ICF, which was developed by the World Health Organization, is the most current model framework. The ICF model describes the effects of a health condition on function (i.e., the ability of the person to work or operate in a proper or particular way) as a multidimensional combination of body function & structures, activities, participation, and contextual factors related to the individual and the environment.

To restate these dimensions into more clinician friendly terminology, body function & structures is related to physical limitations such as lack of range of

motion, joint effusion, or bone fractures; Activities relates to an individual's capacity to perform activities of daily living; Participation provides context for the activities being performed; Lastly, contextual factors refers to the inter-individual differences between patients relating to social and personal factors (e.g., occupation, family history, injury history).

Knee-Related Patient-Reported Outcomes Instruments

There are currently 24 unique PRO instruments developed to assess knee-related function (Wang, Jones, Khair, & Miniaci, 2010). Many of these instruments, however, were not intended to be used in the sports setting. Athletes typically have much greater functional capacity compared to non-athletic adults. Because of this instruments designed for the general public may not function well in athletes.

An independent research group established by the Academy of Orthopaedic Surgeons (AAOS), designed to evaluate PRO instruments, suggested that the KOOS and the International Knee Documentation Committee Subjective Knee Form (IKDC-SKF) were both appropriate for use with sport populations (American Academy of Orthopaedic Surgeons, 2016). The KOOS was recommended for use with Osteoarthritis-related injuries; while the IKDC-SKF, combined with the Marx Activity Rating Scale (Marx, Stump, Jones, Wickiewicz, & Warren, 2001) was recommended for use with ACL-related injuries (American Academy of Orthopaedic Surgeons, 2016).

Limitations of Knee Injury Patient-Reported Outcomes. While the effectiveness of the KOOS and IKDC-SKF have been well documented in the general population (Irrgang et al., 2001; Paxton, Fithian, Stone, & Silva, 2003; Roos, Roos, Lohmander, Ekdahl, & Beynnon, 1998) there has been limited research investigating these instruments in athletic and other high functioning populations. Furthermore, there are additional concerns related to ceiling effects, and assessment of environmental factors related to knee injuries, which limit the use of these instruments in clinical practice. First, both instruments suffer from significant ceiling effects (Fries, Rose, & Krishnan, 2011; Roos & Toksvig-Larsen, 2003). Ceiling effects are problematic because they limit the capability of instruments to measure the upper ranges of functional ability. This is not a significant issue in the early stages of treatment, or in low functioning populations. In athletes, however, where functional capacity is higher than the general population (Muaidi, Nicholson, & Refshauge, 2009), the instruments will become less useful over time as the athlete recovers. Secondly, no theoretical models were used to guide the development of these instruments. As a result, these instruments are not capable of evaluating all of the domains of function as described by the ICF. Specifically, there are virtually no items relating to environmental function, which can influence how a patient perceives the progress of their treatment.

Item Response Theory

Item Response Theory (IRT) is a collection of advanced statistical models that can be used to improve knee-related PRO. Previous studies, involving PRO

instruments have used IRT to identify potential ceiling effects, and to develop new items which effectively minimized the ceiling effects (Fries et al., 2011).

These models are valuable because they allow for separation of item parameter and person ability estimates. The separation of parameters is important because it allows for the development and improvement of these instruments. Evaluation of the KOOS, IKDC-SKF, and Marx Activity Scale with IRT models will allow for identification of ranges in functional ability that are currently poorly assessed in athletic populations. This information can be used to guide the development of new items, which will improve overall assessment of knee function.

Needs Statement

Establishment of best treatment practices requires documentation of PRO throughout the entire recovery process. Limitations of the current instruments restrict their ability to evaluate function, as defined by the WHO, throughout the recovery process, specifically relating to insufficient evaluation of environmental factors, lack of sufficient evidence supporting use in athletic populations, and ceiling effects. These issues must be addressed to improve the current instruments and make them more clinically relevant. Development of new items that [1] evaluate environmental factors related to function, [2] minimize ceiling effects, and [3] perform properly in athletic populations are desperately needed.

CHAPTER II

LITERATURE REVIEW

Knee pain is one of the most common musculoskeletal conditions with recurrent episodes in 6% of the population over 30 and 25% of the population over 55 (Hunter & Felson, 2006). In the United Kingdom, knee pain is the second most common justification for primary-care visits related to musculoskeletal conditions ranging from 16-20% (Urwin et al., 1998; Webb et al., 2004). In the United States, from 1999 to 2008, 6.6 million knee-related injuries were identified as the primary cause of emergency-room visits (Gage, McIlvain, Collins, Fields, & Comstock, 2012). Knee-related injuries were highest among young adults and adolescents (Gage et al., 2012): with participation in recreation and sport being the leading cause of knee-related injuries; accounting for 49.3% of all knee-related injuries (Gage et al., 2012).

Sports Knee Injury Epidemiology

The National Collegiate Athletic Association (NCAA) Injury Surveillance Program (ISP) began in 1982 to provide current and reliable data on sports-injury trends in intercollegiate athletics (National Collegiate Athletic Association Sports Science Institute). Since 2009, the Datalys Center for Sports Injury Research and Prevention has managed the NCAA ISP. Today, Datalys collects and reports injury statistics on 25 different NCAA sports from all divisions. The recent 2012 NCAA football kick-off rule change is an example of the NCAA ISP data put to use, which showed a disproportionate number of injuries occurring during kickoffs (National Collegiate Athletic Association).

In intercollegiate sports, knee-related injuries are among the most common injuries. The prevalence of knee injuries across 14 intercollegiate sports are displayed in table 1. From 1988-2004, 178,925 injuries occurred across 14 intercollegiate men's and women's sports (Agel, Dick, et al., 2007; Agel, Dompier, et al., 2007; Agel, Evans, et al., 2007; Agel, Olson, et al., 2007; Agel, Palmieri-Smith, et al., 2007; Agel, Ransone, et al., 2007; Dick, Ferrara, et al., 2007; Dick, Hertel, et al., 2007; Dick, Lincoln, et al., 2007; Dick, Putukian, et al., 2007; Dick, Romani, et al., 2007; Dick, Sauers, et al., 2007; S. W. Marshall, Covassin, et al., 2007; S. W. Marshall, Hamstra-Wright, et al., 2007). Among these injuries, 14.5% were knee related (Agel, Dick, et al., 2007; Agel, Dompier, et al., 2007; Agel, Evans, et al., 2007; Agel, Olson, et al., 2007; Agel, Palmieri-Smith, et al., 2007; Agel, Ransone, et al., 2007; Dick, Ferrara, et al., 2007; Dick, Hertel, et al., 2007; Dick, Lincoln, et al., 2007; Dick, Putukian, et al., 2007; Dick, Romani, et al., 2007; Dick, Sauers, et al., 2007; S. W. Marshall, Covassin, et al., 2007; S. W. Marshall, Hamstra-Wright, et al., 2007). The prevalence of knee injuries varies across sports with the highest prevalence of knee-related injuries occurring in contact sports (e.g., wrestling, football, basketball).

Table 1. Prevalence of Knee-Related Injuries for Each Sport

Sport	Knee-Related Injuries	Total Injuries	Percentage
Men			
Baseball	337	8,346	4.0%
Football	13,968	84,095	16.6%
Basketball	1,252	12,044	10.4%
Ice Hockey	986	6,639	14.9%
Lacrosse	584	4,845	12.1%
Soccer	1,652	12,974	12.7%
Wrestling	1,774	9,723	18.2%
Women			
Basketball	1,620	10,211	15.9%
Ice Hockey*	50	431	11.6%
Lacrosse	405	3,392	11.9%
Soccer	1,740	11,209	15.5%
Softball	537	5,336	10.0%
Volleyball	719	6,941	10.4%
Gymnastics	367	2,739	13.4%

Note. The above table displays the knee-related injury rates from 14 intercollegiate sports as reported by the NCAA Injury Surveillance Program in 2007 (Agel, Dick, et al., 2007; Agel, Dompier, et al., 2007; Agel, Evans, et al., 2007; Agel, Olson, et al., 2007; Agel, Palmieri-Smith, et al., 2007; Agel, Ransone, et al., 2007; Dick, Ferrara, et al., 2007; Dick, Hertel, et al., 2007; Dick, Lincoln, et al., 2007; Dick, Putukian, et al., 2007; Dick, Romani, et al., 2007; Dick, Sauers, et al., 2007; S. W. Marshall, Covassin, et al., 2007; S. W. Marshall, Hamstra-Wright, et al., 2007).

*Injury Surveillance Program Data for Women's Ice Hockey is only reported for 2000-2004.

Consequences of Knee Injuries

Knee-related injuries can be especially problematic for athletes because of the potential for long-term physiological effects. Crude estimates suggest that 10-20 years after an anterior cruciate ligament (ACL) or a meniscal tear, one out of

every two patients will develop osteoarthritis, often demonstrating significant pain, functional limitations and diminished quality of life (L. S. Lohmander et al., 2007). Athletes with a history of knee-related injury often report lower health-related quality of life compared to those without history of knee-related injuries (Lam et al., 2017).

Extended Lost Time from Competition

Majority of injuries occurring in sports are relatively mild and require minimal lost time from sports (Agel, Dick, et al., 2007; Agel, Dompier, et al., 2007; Agel, Evans, et al., 2007; Agel, Olson, et al., 2007; Agel, Palmieri-Smith, et al., 2007; Agel, Ransone, et al., 2007; Dick, Ferrara, et al., 2007; Dick, Hertel, et al., 2007; Dick, Lincoln, et al., 2007; Dick, Putukian, et al., 2007; Dick, Romani, et al., 2007; Dick, Sauers, et al., 2007; S. W. Marshall, Covassin, et al., 2007; S. W. Marshall, Hamstra-Wright, et al., 2007). In some cases however, injuries can be more severe, requiring extended lost time (i.e., 10 or more days) from sports competition. Knee-related injuries are one of the leading causes of extended lost time in sports competition (Agel, Dick, et al., 2007; Agel, Dompier, et al., 2007; Agel, Evans, et al., 2007; Agel, Olson, et al., 2007; Agel, Palmieri-Smith, et al., 2007; Agel, Ransone, et al., 2007; Dick, Ferrara, et al., 2007; Dick, Hertel, et al., 2007; Dick, Lincoln, et al., 2007; Dick, Putukian, et al., 2007; Dick, Romani, et al., 2007; Dick, Sauers, et al., 2007; S. W. Marshall, Covassin, et al., 2007; S. W. Marshall, Hamstra-Wright, et al., 2007). The prevalence of extended time lost injuries related to the knee are displayed in table 2. In football, 30.6% of extended lost time injuries were knee-related (Dick, Ferrara, et al., 2007). Similar

results have been reported across other sports with prevalence of 18.6% and 24.7% for volleyball (Agel, Palmieri-Smith, et al., 2007), and women's lacrosse (Dick, Lincoln, et al., 2007), respectively.

Table 2. Prevalence of Extended Lost Time Knee-Related Injuries

Sport	Knee-Related Injuries	Extended Time-Loss Injuries	Percentage
Men			
Baseball	143	2,102	6.8%
Football	6,913	22,556	30.6%
Basketball	415	2,187	19.0%
Ice Hockey	418	1,740	24.0%
Lacrosse	233	1,007	23.1%
Soccer	586	2,173	27.0%
Wrestling	771	2,933	26.3%
Women			
Basketball	786	2,468	31.8%
Ice Hockey*	-	-	-
Lacrosse	195	789	24.7%
Soccer	763	2,136	35.7%
Softball	235	1,248	18.8%
Volleyball	262	1,406	18.6%
Gymnastics	194	916	21.2%

Note. Note. The above table displays the knee-related injury rates from 14 intercollegiate sports as reported by the NCAA Injury Surveillance Program in 2007 (Agel, Dick, et al., 2007; Agel, Dompier, et al., 2007; Agel, Evans, et al., 2007; Agel, Olson, et al., 2007; Agel, Palmieri-Smith, et al., 2007; Agel, Ransone, et al., 2007; Dick, Ferrara, et al., 2007; Dick, Hertel, et al., 2007; Dick, Lincoln, et al., 2007; Dick, Putukian, et al., 2007; Dick, Romani, et al., 2007; Dick, Sauers, et al., 2007; S. W. Marshall, Covassin, et al., 2007; S. W. Marshall, Hamstra-Wright, et al., 2007).

*The NCAA Injury Surveillance Program reports did not include data on extended lost time injuries for Women's Ice Hockey.

Consequences of Extended Time Loss Injuries

Sports-related injuries resulting in extended lost time can be a traumatic event for athletes. In some cases, specifically with older athletes, there is a risk that severe injuries can end a playing career. In 1990 injury surveillance of an elite-Swedish soccer team indicated that only 4 out of 12 players who sustained an ACL rupture returned to play at an elite level 18.5 months after injury; the others transferred to lower divisions or were still in rehabilitation (Engström, Forssblad, Johansson, & Tornkvist, 1990). In the National Basketball Association (NBA) 22% of players who required surgical interventions to repair ACL ruptures, over 10 seasons, and 19.4% of players with isolated meniscal tears over 21 seasons, did not return to a sanctioned NBA game (Busfield, Kharrazi, Starkey, Lombardo, & Seegmiller, 2009; Yeh, Starkey, Lombardo, Vitti, & Kharrazi, 2012).

A major concern that discourages many athletes from returning to competition is decreased level of performance following injury. Evaluation of playing performance in Australian-football athletes indicated that following injury, player performance was significantly lower in the two games after returning to competition compared to playing performance in the two games prior to injury (Verrall, Kalairajah, Slavotinek, & Spriggins, 2006). In this study however, the pre-injury performance levels were equivalent to the performance levels of athletes across the entire season (Verrall et al., 2006); suggesting that the athletes may have been returned to competition too soon following injury. Appropriate documentation of injuries following injury may help to minimize returning players to competition too soon.

Injury Evaluation Models

Documentation of patient outcomes is necessary to ensure that athletes have sufficiently recovered prior to returning to sports competition following injury. A variety of injury evaluation models have been developed over the years to monitor changes in patient function following injury. The following sections will provide a brief overview of injury evaluation models, which includes early disablement models as well as the current functional model proposed by the World Health Organization (WHO).

Disablement Models

Nagi Model. One of the earliest known disablement models was developed by an Egyptian sociologist, Saad Nagi, who believed that medicine had a restricted view of disability (Masala & Petretto, 2008). Nagi believed that disablement was a reflection of physical or mental limitations in a social context; creating a gap between the individual's capabilities and the demands created by the physical and social environment (S. Nagi, 1965).

Nagi's disablement model included four concepts: active pathology, impairment, functional limitation, and disability (figure 1) (S. Nagi, 1965). These four concepts refer to the specific disease condition; the anatomical changes resulting from this condition (e.g., abnormality at the tissue, organ, or body system level); limitations in a person's ability to perform the tasks associated with daily living; and limitations in performing socially defined roles, respectively (S. Nagi, 1965).

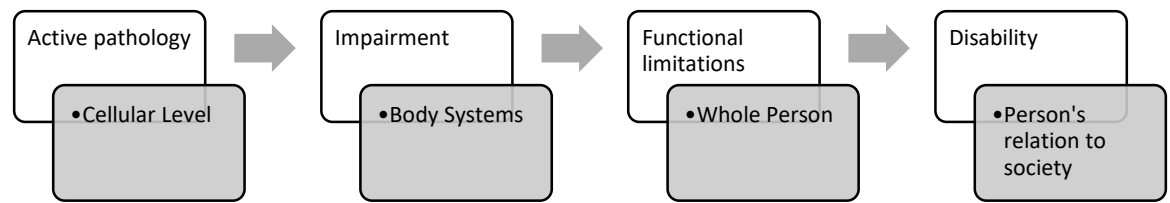


Figure 1. Nagi's Disablement Model. Adapted from Nagi, 1965 (S. Nagi, 1965).

Nagi's disablement model helped to revolutionize the role of the environment and society in disability and health evaluation models. Conceptually the Nagi model was used to explain the process and underlying mechanisms which diseases and injuries impacts a person's ability to function (i.e., perform their expected role in society). The Nagi model placed greater emphasis on the consequences of disease, rather than the disease itself.

This changing paradigm led to a reflection of the relationship between pathologies and disease consequences. Increasing awareness of the consequences of disease led to a large epidemiological study in Great Britain where a growing number of individuals were living in extreme poverty. This study, led by Amelia Harris, was targeted at detecting the correlation between poverty and reduced autonomy, and also at developing tools for interventions and prevention (Harris, 1971).

International Classification of Impairments, Disabilities, and

Handicaps. Harris' research was criticized due to methodological concerns which some believed under-represented the true number of "disabled", "impaired", and "handicapped" individuals living in Great Britain (Masala & Petretto, 2008). A small research group coordinated by Phillip Wood and Elizabeth Badley tried to clear up some of the confusions in Harris' work; for example, Harris often used the terms "handicapped" and "impairment" interchangeably which caused confusions (Badley, Thompson, & Wood, 1978). In addition to investigating Harris' work, Wood was also serving as a consultant for the World Health Organization. His experiences with Harris' work, led to the development of the International Classification of Impairments, Disabilities, and Handicaps (ICIDH) model; which was presented to the public at the 29th Assembly of the WHO (World Health Organization, 1980).

The WHO ICIDH included three different levels of pathological disease consequences, relating to different levels of experience and of individual awareness (figure 2). Although the original model was intended as more of a framework with interconnected disease consequences, the linear representation of the figure led some to misinterpret and criticize the original model (World Health Organization, 2001). Despite criticisms the ICIDH model was the first internationally shared conceptual health framework, which was translated into 13 languages (World Health Organization, 1980).

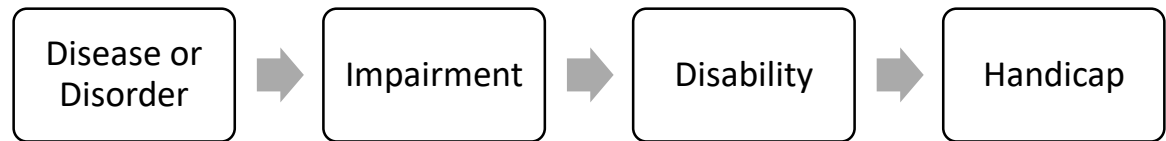


Figure 2. International Classification of Impairments, Disabilities, and Handicaps. Adapted from World Health Organization, 1980 (World Health Organization, 1980).

Functional Models

Dissatisfaction with the current disablement models arose from individuals who, because of the impairments and autonomy limitations, were the recipients of discrimination and actions that limited their freedoms, as well as human and social rights (Masala & Petretto, 2008). Disablement models were developed from a need to better classify individuals, specifically those in need of economic and welfare support measures. The Fundamental Principles of Disability, an activist-rights group proposed that “physical disability is a form of social oppression” and that “society disables people with impairments” (Union of the Physically Impaired Against Segregation, 1975).

A limitation of the ICIDH and other disablement models is their inability to separate social limitations and causes of these limitations (Guccione, 1991). In addition, there is considerable overlap in the disability and handicap dimensions (S. Z. Nagi, 1991), which make classification of individuals challenging when

limitations are multidimensional (Jette, 1994). These limitations as well as criticisms against use of the terms “disabled” and “handicapped”, which had negative connotations led to the revision of disablement models to the more current functional models; which emphasize a more broad and modern view of health and disability.

International Classification of Function, Disability, and Health. The International Classification of Function, Disability, and Health (ICF) was proposed at the 54th World Health Assembly as an improvement of previously used disablement models (World Health Organization, 2001). Contextual factors related to the individual and the environment were included. In addition, the focus of the model transitioned to one that evaluates overall health rather than the consequences of disease. Furthermore, complaints about the linear structure of disablement models led to an inter-connected framework where each dimension contributes to an individual’s level of function and health. Limitations in one dimension can influence other dimensions. For example, a patient who lives in the fifth floor of a large apartment building may have greater difficulty engaging in activity than one who lives in a ranch-style home in the suburbs. The decreased activity could prolong their recovery time or result in increased levels of pain.

The multi-dimensional ICF model includes body functions & structures, activities, participation, environmental factors and personal factors (figure 3).

Body function and structures refer to the physiological functions of body systems and the anatomical parts of the body, such as limbs and their components. Irregularities in function/structure are referred to as impairments,

which are defined by a loss in range of motion, muscle weakness, and pain and fatigue (Snyder et al., 2008; World Health Organization, 2001).

Activities represent the act of completing a task and is related to a patient's perceptions of function. Participation refers to the involvement of a patient in real world situations. Limitations in activity and participation may be demonstrated through mobility dysfunctions or difficulty with walking, climbing steps, carrying or grasping objects (Snyder et al., 2008; World Health Organization, 2001).

Environmental factors, describe the external influences affecting health status of a patient. Environmental factors may be physical, social, cultural, or institutional in nature (P. Rosenbaum & Stewart, 2004; Snyder et al., 2008). Personal factors describe the history of the patient and individual differences that may influence health or function. The separation of environmental factors from activity limitations and participation is a major advantage of the ICF over traditional disablement models; which allows for a more in-depth investigation of the factors influencing function and health.

The ICF framework has been used in a variety of settings including the Multi-Country Survey Study and the World Health Survey Program to measure health status of the general population in 71 countries (World Health Organization, 2016). The ICF model framework has also been receiving increased attention in athletic settings due to the increased emphases on evidence-based practice. In 2008, the Journal of Athletic Training published a series of papers urging athletic trainers to adopt the ICF as a framework for

assessing clinical outcomes in athletic training (McLeod et al., 2008; Snyder et al., 2008). Functional health models help to standardize language across professions leading to better consistency and improved quality of care for patients.

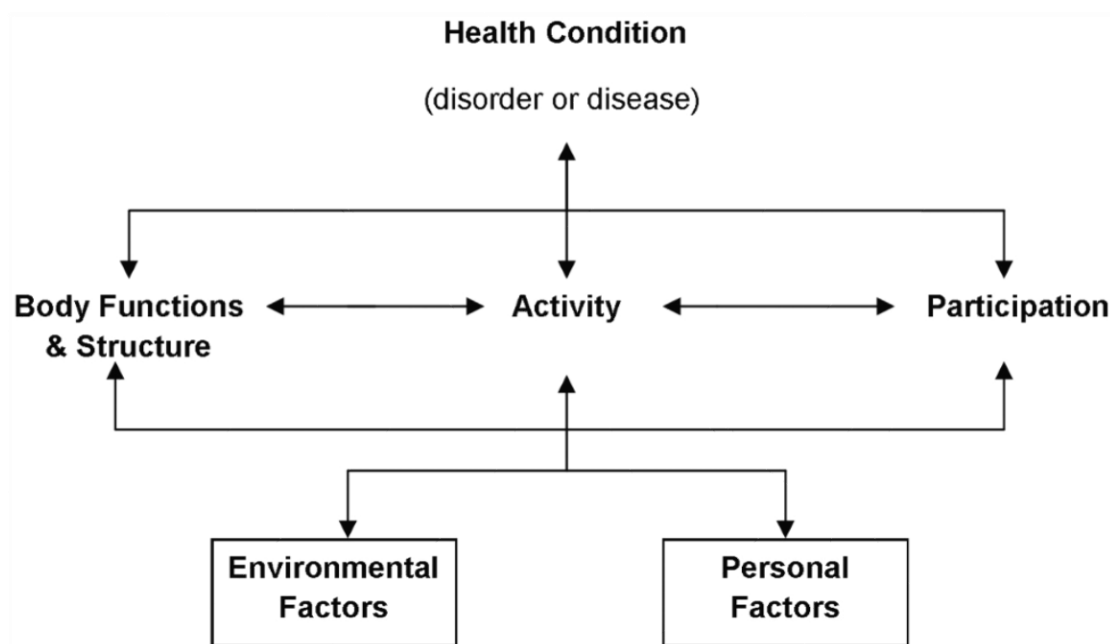


Figure 3. World Health Organization International Classification of Functioning, Disability, and Health. Adapted from The World Health Organization International Classification of Functioning, Disability, and Health (World Health Organization, 2001).

Historical Review of Knee-Related Injury Outcomes Measures

Many different instruments have been developed to assist physicians with documentation of knee outcomes. Early outcome instruments for the assessment of knee-related function were developed by physicians in order to measure and compare the effectiveness of surgical techniques following knee ligament surgery. These early instruments were often developed with no evidence of validation beyond the expertise of the development team. Some of these instruments are still in use today despite changes in evaluation models and surgical techniques. The following section provides a brief review of some of the more influential instruments as well as their advantages and disadvantages for use.

Larson Scoring Scale

The original Larson scoring scale represents one of the earliest and first used tools for evaluating surgical interventions of the knee (Larson, 1974). The scale consisted of 15 items, which primarily assessed stability of the knee joint following ligament reconstruction. This instrument was used by surgeons to demonstrate efficiency of surgical techniques (Oretorp, Gillquist, & Liljedahl, 1979; Solonen & Rokkanen, 1967). No documentation has been provided on the development or validation of this instrument. The key success for this instrument is that it introduced surgeons to the idea of quantifying results in an effort to demonstrate and measure surgical success.

Marshall System

The first standardized evaluation system for knee ligament injuries was developed by Marshall et al., 1977. The Marshall system was completed by the physician and assessed symptomology and functional disability for patients after knee-ligament surgery. Patient scores ranged from 0 to 50; where scores between 41-50 indicated normal or near normal function; and scores less than 30 indicating severely disabled, with marked signs and symptoms of a compromised knee (J. L. Marshall et al., 1977).

While the Marshall system was short and easy to read, it had a major shortcoming. Despite experiencing significant functional deficiencies, some patients would still receive high scores (F. Hefti, Gächter, Jenny, & Morscher, 1982). This over-rating of patients severely hindered the use of this instrument in the clinical setting for documentation of patient outcomes. Furthermore, there have been no published studies examining the validity or reliability of this instrument (Lysholm & Tegner, 2007).

Lysholm-Tegner Rating System

In 1982, Lysholm and Gillquist created the Lysholm score for assessing knee ligament surgery follow-up (Lysholm & Gillquist, 1982). The Lysholm score was adapted from the original Larson scoring scale. Four of the original items, which include: genu varum or valgum, genu recurvatum, flexion contracture, and patellar abnormality were removed. An additional item “instability”, defined as ‘giving away’, was added. The Lysholm score ranges from 0 to 100-points where

a score of 100 indicates a normal knee. Scores are determined by summing the ratings from each category (e.g., function, activities, squatting, walking, etc.).

The Lysholm score was later revised to include only subjective items and renamed The Lysholm-Tegner rating system (Tegner & Lysholm, 1985). The scoring system for the revised instrument remained the same (e.g., 100-point scale). Instability and pain were each attributed 25-points; with a total of 8 items. In addition, an activity-scale question was added, which asked patients to rate the highest level of activity they can currently perform. Options for the question range from 0 (sick leave or disability pension because of knee problems) to 10 (competitive sports- soccer, football, rugby [national elite]).

Development of the Tegner-activity scale was important because it was one of the first knee-related instrument whose purpose was not related to comparing surgical techniques. The Tegner-activity scale was used to compare pre-operative and post-operative activity levels (Tegner & Lysholm, 1985). The individual-centered evaluation model was an important advancement for knee assessment which helped to spark the development of more in-depth tools.

Although this instrument was developed over three decades ago, the Lysholm-Tegner rating scale is still used today (Briggs et al., 2009). There are some limitations however, that should be acknowledged which limit the clinical usefulness of this instrument.

There is currently no “gold-standard” instrument for assessment of knee function, which makes validation of these instruments a challenging task. A recent validation study suggested that although the Lysholm score was positively

correlated with modern evaluation tools ($\rho = .78$; $p < .01$) the Tegner-activity scale was poorly correlated with modern evaluation tools ($\rho = .22$; $p < .01$).

Validation of instruments is important to ensure they are measuring the construct they are intended to measure. The low correlation between the Tegner-activity scale and modern knee-evaluation instruments suggest that the Tegner-activity scale may not be a good measure of knee-related function.

Content validity for the Tegner-activity scale is questionable. Categories for the Tegner-activity scale are assigned arbitrarily with no evidence to support the hierarchy of activities. Moreover, some of the categories have a large degree of overlap. For example, “Work-light labor” is listed for both levels 2 and 3.

Jogging at least 5 times per week is listed at a lower activity than running; however, no time is specified for running. Therefore, a patient who runs once for five minutes one time per week would be classified higher than a patient who jogs consistently five or six days per week.

Another major limitation of this instrument is that no rational or theory has been provided for the development of the scoring scale, or the weighting of scores across items in each domain of the Lysholm score. Pain and instability account for 50% of the scoring while squatting and stair climbing, questions related to activity limitations account for a combined 15%. This is concerning because a patient with a high pain threshold, and minimal instability could obtain a relatively high score despite significant limitations in activity and mobility. On the other hand, a patient with no functional limitations, yet experiencing moderate to high pain could receive a low overall score indicating poor knee function. The

validity of this scoring system is questionable because it is difficult to assess what is actually being assessed. In addition, 68% and 61% of patients scored the maximum score for Support and Locking domains of the Lysholm score, respectively; while 8% of patients responded in the lowest category of the Tegner-activity scale.

Lukianov Anterior Cruciate Ligament Evaluation Format

Lukianov et al., proposed a synthesis of all known knee evaluation systems (Lukianov, Gillquist, Grana, & Dehaven, 1987). This documentation system was criticized for its length (i.e., 15-pages) which was impractical in the clinical setting where time is a limiting factor. Because of the extended length of this instrument it was not popular in the field and no studies were published examining the validity of this instrument. Furthermore, because this instrument is only a compilation of previously developed instruments, the limitations of the previous instruments were included in this instrument as well.

Cincinnati Knee Rating Scale

The Cincinnati Knee Rating Scale (CKRS) was developed in 1982 by Noyes et al., to assess efficiency of rehabilitation and activity modifications in patients with ACL-deficient knees (F. Noyes, Matthews, Mooar, & Grood, 1983; F. R. Noyes, Mooar, Matthews, & Butler, 1983). The original CKRS included six functional components (symptoms, pain location, swelling, giving way, other symptoms, and overall activity level) that assessed patient symptoms at various levels of activity and in activities of daily living.

The CKRS has been revised a number of times since its initial development. The current version includes seven functional components: subjective assessment, patient history, knee examination, objective testing, operative procedures and articular cartilage rating, post-operative complications, and overall rating (Barber & Noyes).

A unique aspect of the CKRS is the inclusion of both subjective patient ratings and physician documented clinical-based outcomes. Authors of the instrument claim that while subjective assessment of symptoms and functional limitations are important, the final outcome of a specific treatment must also take into account objective measures such as physical findings, radiographs, and arthrometer (i.e., joint angle/ goniometer) measurements (Barber & Noyes).

While these data are helpful for determining the efficiency of treatment outcomes in the clinical setting (Swiontkowski, 2005), these outcomes may be misinterpreted and inappropriately used to describe function (Binkley, 1999). For example, a patient engaging in rehabilitation following treatment may demonstrate an increase in knee flexion from 70° to 120° that suggests improvement. This patient however, may still be unable to perform activities necessary for daily living or engage in sports participation. Knee joint-angle measurements in 684 patients at 12-months following total knee arthroplasty found no correlations between maximum knee flexion and function when evaluated through patient-reported outcomes (PRO) (Miner, Lingard, Wright, Sledge, & Katz, 2003).

The complexity and time required to complete the CKRS make this instrument impractical for everyday use; specifically, in athletic training settings, where radiographs may not be available. In general, the domains included in this instrument closely align with the model proposed by the WHO ICF. Noyes et al., argued that many of the instruments developed around the same time were flawed because they failed to account for the individual differences in patients, including lifestyle changes, and behaviors that influence patient function (F. R. Noyes, Barber, & Mooar, 1989).

International Knee Documentation Committee

The limitations of currently used instruments, as well as inconsistencies in terminology among clinicians, led to the development of the International Knee Documentation Committee (IKDC) in 1987 (E. Hefti, Müller, Jakob, & Stäubli, 1993). The committee consisted of knee surgeons from Europe and America with the goal of developing a common language and a set of minimum standards for the evaluation of knee ligament surgeries. The proposed guidelines for the “ideal” evaluation and documentation systems are outlined in figure 4.

- ❖ The form should be brief (not more than one page)
- ❖ It should cover all relevant findings.
- ❖ It should be easy to fill out without a long introduction.
- ❖ Arbitrary judgement should be minimized.
- ❖ In the final rating emotional terms (e.g., “excellent”, “good” or “poor” should not be used.
- ❖ It should be impossible for a knee to be rated “nearly normal” when a serious problem exists. The problem should be apparent.
- ❖ Numerical scores should not be used.
- ❖ The patient’s preinjury activity level should be taken into account.
- ❖ The system should be computer-compatible.

Figure 4. International Knee Documentation Committee “Ideal” Evaluation and Documentation Guidelines (E. Hefti et al., 1993).

Using the guidelines, the IKDC developed the IKDC’s standard evaluation form. The standard evaluation form was approximately one page in length and included the following sections: symptoms, range of motion, ligament examination, compartmental findings, harvest site pathology, x-ray findings (degenerative joint disease), functional tests, patient subjective assessment, and basic demographic information.

One major difference between the standard evaluation form and previous outcomes instruments was the lack of a numerical scoring system. Early instruments often utilized a 100-point (0 to 100) composite scoring system where total knee function was determined from the sum of scores in each assessment category. These original scoring systems were designed in an effort to demonstrate the “quality” of various surgical techniques so that surgeons could recommend their services over another surgeon. As surgical procedures became more standardized (Bach & Boonos, 2001; Han, Seong, Lee, & Lee, 2008; Kurosaka, Yoshiya, & Andrish, 1987) the need to demonstrate the efficiency of one technique over another has diminished. Thus greater emphasis on evaluating rehabilitation and function are required.

Interestingly most instruments utilized the same classification thresholds (e.g., 90-100 points = Excellent, 80-89 points = Good, 70-79 = Fair, < 70 = Poor). These classification thresholds have not been properly validated. Under the 100-point scale system the difference between “poor” knee and a “fair” one is one point. This limitation was also acknowledged by the IKDC, which felt that a more subjective categorical classification system would be beneficial (e.g., normal function, nearly normal, abnormal, severely abnormal) (E. Hefti et al., 1993). Overall knee function is classified based upon the lowest classification from each functional domain. The IKDC selecting this scoring system on the premise that a “surgically repaired knee could not result in function exceeding the non-injured limb”, and that “a knee with abnormal findings within any category should not be classified as a *normal* knee” (E. Hefti et al., 1993).

Although there are seven functional domains included in the standard evaluation form (effusion, passive motion deficit, ligament examination, compartment findings, harvest site pathology, x-ray findings, and functional tests) only three of the seven domains, effusion, passive motion deficit, and compartment findings are used to determine overall knee function score. Using the new subjective scoring method, a patient with no swelling or functional limitations, with crepitus (crackling, crinkly, or grating feeling or sound in the joint), and mild pain would be described as severely abnormal; regardless of classification in the other two categories. Pain, is a complex phenomenon however and can vary greatly from individual to individual (Coghill, McHaffie, & Yen, 2003). In addition, the presence of crepitus, does not directly indicate the presence of a pathology or injury (Fulkerson, 2004). Though the instrument is reported to assess knee function, this statement is questionable considering the lack of validity supporting this new scaling system. Furthermore, inclusion of assessments that have no impact on the patients score seems unnecessary and provides no useful information regarding treatment outcomes.

Patient-Reported Outcomes

The Patient Protection and Affordable Care Act of 2010 was enacted to increase quality and accessibility of health care options in the United States ("The Patient Protection and Affordable Care Act," 2010) and resulted in the development of the patient-centered outcomes research institute (PCORI) (Selby, Beal, & Frank, 2012). Development of more patient-focused health

outcomes has been a key focus of many health organizations such as PCORI and the National Institute of Health (Ader, 2007; Cella et al., 2007).

The role of the patient in the patient-clinician relationship has evolved due to the variety of health care options currently available. Many traditional outcomes instruments such as the IKDC standard evaluation form, and CKRS produce outcomes that, while valuable to the clinician, are difficult for patients to understand and use effectively to assist with their healthcare decisions. As consumers of those services, patients must decide which options might be right considering their circumstances, preferences, families and cultural beliefs. This changing healthcare structure has made it increasingly important for the development of new instruments that emphasize outcomes that are both relevant and meaningful to the consumer of those services (i.e., the patients); allowing them to make the right decision concerning their own medical treatment (Annas, 1975; Roter, 1977).

Patient-reported outcomes are typically obtained through self-report instruments and include information related to a patient's physical, mental, emotional, and social well-being as described by the ICF model framework. Using PRO, patients, their families, and clinicians, can obtain an understanding of patient function using terminology and outcomes that are more relevant to the family.

Advantages of Patient-Reported Outcomes

Similar to clinician-based outcomes, clinicians use PRO to monitor recovery following injury and provide information about the effects of a treatment

(Greenhalgh & Meadows, 1999). When combined with clinician-based outcome measures, PRO improve patient satisfaction and increase compliance with rehabilitation protocols leading to improved quality of care (Roter, 1977).

When patients are more knowledgeable about treatment outcomes they are better equipped to take a more active role in the medical decision-making process (Greenfield, Kaplan, & Ware, 1985). Because PRO are self-reported they can be completed by the patient prior to meeting with the clinician. Thus, allowing more time for patient-clinician interaction. In addition, PRO are easier for patients to understand, which may help to facilitate communication between patients and clinicians who often value different measures of treatment success (Rothwell, McDowell, Wong, & Dorman, 1997).

Patient-reported outcomes collected using standardized instruments may assist with the recognition of health conditions in the clinical setting that might otherwise be overlooked. It is important however, not to completely disregard clinician-based outcomes. Correlations between clinician-based outcomes and PRO are poor, suggesting that these instruments are measuring different constructs (Kantz et al., 1992; Maly, Costigan, & Olney, 2006).

Clinician-based outcomes such as joint laxity only represent a single dimension of function (i.e., body functions and structures). For example, in a sample of 527 patients following ACL surgery the Lachman's test, a commonly used assessment technique to assess knee-joint laxity and diagnose ruptures of the ACL, was weakly associated ($r = -.42$) with commonly used PRO measures (Sernert et al., 1999). On the other hand, PRO assess limitations in participation

and activities of daily living. PRO also allow for the investigation of environmental and contextual factors such as lifestyle changes, social support system, and role functioning (i.e., occupational/job-related limitations) which are not included in clinician-based outcomes. By including both types of outcomes in the assessment process, a more detailed evaluation of function is possible; which encompasses the environmental and personal factors influencing limitations in activity and participation; as well as the anatomical and structural limitations associated with a specific health condition.

Types of Patient-Reported Outcomes Instruments

A variety of PRO instruments have been developed to assess patient function. The two main classifications of function instruments include generic and specific. Specific-type instruments can be further classified as population-specific or disease-specific (injury/body-region). The advantages and disadvantages of these instrument types have been summarized in table 3.

Table 3. Summary of the Types of Patient Reported Outcomes Instruments

Type of Patient-Reported Outcome Instrument	Advantages	Disadvantages	Examples
Generic	• Broad assessment of health • Useful for comparing health outcomes across a variety of patient populations	• 10-20% less precision than specific type instruments(Leong et al., 2005; Marra et al., 2005; McHorney, Ware Jr, Rogers, Raczek, & Lu, 1992) • Limited diagnostic utility	Short Form 36; Short Form 12
Single Item, Summary	• Minimal patient/clinician burden • Can easily be added to a longer instrument	• Limited scope • Less precise than longer alternative instruments	Global Rating of Change Scale (GROC); Patient Acceptable Symptom State (PASS)
Disease-specific, Body-region specific, Injury-specific	• More precise estimate of health compared with generic instruments(Marra et al., 2005) • High responsiveness for specific health condition • Easily defined measurement objective	• May not assess all dimensions of function equally • Limited for use with specific disease or population • May miss unforeseen side effects not specific to disease condition	Disabilities of the Arm, Shoulder, and Hand (DASH); Knee Injury and Osteoarthritis Outcomes Score (KOOS)
Population-specific	• Highly relevant to population of interest	• Items may not be relevant to populations other than originally intended	Child Health and Illness Profile;

Patient-Reported Outcomes Instruments for the Assessment of Knee-Related Function

Wang et al., identified 24 unique PRO instruments for the assessment of knee-related function (Wang et al., 2010). Among these 24 instruments, three were used to evaluate knee-related function associated with sports injuries in young and middle-aged adults. These instruments included the Knee Injury and Osteoarthritis Outcome Score (KOOS), the Academy of Orthopaedic Surgeons (AAOS) Sports Knee Scale, and the IKDC Subjective Knee Form (SKF).

More recently, the Quality Outcomes Data (QOD) work group, a research group created by the AAOS for the purposes of investigating and evaluating PRO data collection tools, provided a list of acceptable PRO instruments, which represented “Orthopaedic Quality Data” (American Academy of Orthopaedic Surgeons, 2016). Interestingly the AAOS Sports Knee Scale was not included on the QOD list. The QOD work group agreed with the recommendations of Wang et al., suggesting that the KOOS and the IKDC-SKF were recommended for evaluation of knee-related function. The KOOS was recommended for use with Osteoarthritis-related injuries; while the IKDC-SKF, combined with the Marx Activity Rating Scale (Marx, Stump, et al., 2001) was recommended for use with ACL-related injuries (American Academy of Orthopaedic Surgeons, 2016).

Evaluation of knee-related function is important because of the high prevalence of knee-related injuries in sports. The following section will describe the development and measurement properties of the KOOS, the Marx Activity

Scale, and the IKDC-SKF. In addition, the limitations of these instruments as well as their relationship to the ICF model will also be discussed.

Knee Injury and Osteoarthritis Outcomes Score

Critical analysis of knee ligament rating systems such as the Lysholm score, the CKRS, and the IKDC standard evaluation form suggested that each instrument was measuring different constructs (Sgaglione, Del Pizzo, Fox, & Friedman, 1995). In addition, these instruments were not self-administered which increased clinician burden when evaluating patient outcomes. Furthermore, these instruments emphasized clinician-based outcomes, which are not very meaningful to patients. In order to address these key issues, Roos et al., developed the KOOS in 1998 to measure knee-related function with a standardized patient-centered measure that could be administered at a low cost (Roos et al., 1998).

Development. A panel of experts consisting of patients, orthopaedic surgeons, and physical therapists from Sweden and the United States were asked to identify symptoms and functional limitations associated with ACL or meniscus injuries. Seven factors were identified by the panel: pain, early disease-specific symptoms, late disease-specific symptoms, function, quality of life, activity level, and satisfaction. A pilot study including patients with a history of meniscal surgery in the past 20 years ($n = 75$) evaluated the factors that were most important to patients during their recovery. Results from the pilot study suggested that pain, swelling, stiffness, and the ability to run, jump, kneel, and squat were of the most importance.

Data from the pilot study, expert panel, and literature review were combined to develop the KOOS. To ensure content validity of the KOOS with older populations the authors included the Western Ontario and MacMasters Universities (WOMAC) Osteoarthritis Index (an outcome measure covering pain, stiffness, and function) (Bellamy, 1988) in the KOOS questionnaire. Satisfaction and activity level factors were excluded from the final version of the instrument due to the inability of the review panel to come to a consensus on wording that would be applicable for all situations.

Finalized versions of the KOOS were developed simultaneously in English and Swedish including 42 items related to pain ($n = 9$), symptoms ($n = 7$), activities of daily living ($n = 17$), function in sport and recreation ($n = 5$), and knee-related quality of life ($n = 4$). The test takes approximately 10 minutes to administer, and requires no formal training or equipment.

Scoring system. All items are scored from 0 to 4. Scores received on each item are summed to obtain scale scores; there is no total score. The range of possible scale scores are as follows: pain 0 – 36; symptoms 0 – 28; activities of daily living 0 – 68; sport and recreation 0 – 20; and knee-related quality of life 0 – 16. Scale scores are transformed to a 0 – 100 scale, by multiplying the raw score by 100, and dividing by the highest possible raw score. For example, a patient with a scale score of 16 for the pain scale will have a total pain score of 66 ($16 * 100 / 36 = 66$).

The scores for each scale range between 0 and 100, with 0 indicating extreme knee problems and 100 indicating no knee problems. The scores for

each of the five subscales are presented as an outcome profile, which collectively represents a patient's level of knee-related function.

Measurement Properties. Reliability for the KOOS was assessed by administering the instrument twice in patients who had sustained an ACL injury prior to surgical intervention. Twenty-one patients were recruited for reliability analysis; however, only 13 patients' data were available for the second testing period. The average time interval between testing sessions was 3.6 ± 2.6 days. Below-average (0.70 – 0.79), average (0.80 – 0.89), and above-average (≥ 0.90) acceptable intraclass correlation coefficients (ICC) were reported for the KOOS scales (pain 0.85; symptoms 0.93; activities of daily living 0.75; sports and recreation function 0.81; and knee-related quality of life 0.86). None of the KOOS scales had unacceptable ICC values (< 0.70).

While the reliability of the KOOS is exceptionally high with this population; the small sample size, lack of confidence intervals for ICC values, and lack of information on ICC model selection is concerning. Meta-analysis of ICC data has indicated that model selection can significantly influence reliability estimates for test scores (Farnsworth II et al., 2017). Furthermore, the reliability analysis was limited to a single injury type (ACL injury). It is unknown whether these results would be consistent for other injury types.

Evidence of validity for the scale scores of the KOOS has been assessed using convergent and known-group difference methods. Spearman's rank correlations were used to assess evidence of convergent validity through associations between scores of the KOOS scales and related subscales of the

SF-36. Coefficients ranged from weak to moderate with the strongest correlations identified between the activities of daily living (KOOS activities of daily living & SF-36 physical function; $r = 0.57$), and pain (KOOS pain & SF-36 bodily pain; $r = 0.46$) scales (Roos et al., 1998). The SF-36 does not contain an equivalent sport and recreation subscale, however, the KOOS function in sport and recreation subscale was moderately correlated with the SF-36 physical function subscale ($r = 0.47$) (Roos et al., 1998).

The coefficient of determination (R^2) is a statistical measure that indicates how much of the variations in a variable can be explained by another variable. The higher this value the more strongly associated the two variables are. The pain scale of the KOOS, and bodily pain subscale of the SF-36 are supposedly measuring the same constructs (i.e., subjective pain). These two scales however, only share 21% common variance. The coefficient of determination is only slightly higher between the activities of daily living scale of the KOOS and the physical function subscale of the SF-36 ($R^2 = 32\%$).

Responsiveness of KOOS scales were assessed through comparison of pre-operative and post-operative scores for patients who had sustained an ACL injury. Significant improvements in scores were noted for pain ($p = 0.02$), activities of daily living ($p < 0.01$), and knee-related quality of life ($p < 0.01$) scales at three-months post-operation (Roos et al., 1998). At six-months post-operation significant improvements were noted for all five KOOS scales compared with pre-operation scores. Effect sizes were computed for six-month change scores: pain = 0.84; symptoms = 0.87; activities of daily living = 0.94;

sports and recreation function = 1.16; knee-related quality of life = 1.65 (Roos et al., 1998). It was not specified whether the effect sizes represent Hedge's G or Cohen's D.

Relationship to ICF. The ICF model was developed as a framework to guide the development of assessment tools related to function. Items of the KOOS can be attributed to many of the dimensions of the ICF.

Body function & structures is assessed through some of the symptom-related questions of the KOOS (e.g., "Does your knee catch or hang up when moving"; "Can you straighten your knee fully"). Activity limitations are assessed through the activity of daily living scale, which includes 17 items related to activities. These items however, are associated with low levels of function. For example, domestic-related activities (e.g., cooking, dusting, scrubbing floors) and ascending/descending stairs) are among the most challenging items. For patients with naturally high levels of function, such as athletes, these items may not be challenging enough to be representative of typical daily activities. Participation is assessed through sport and recreation as well as quality of life scales, which includes nine items. Contextual factors (e.g., environmental and personal factors) are not assessed in the KOOS.

Additional Limitations. Overall, the KOOS has acceptable measurement properties for reliability and responsiveness. Although validation evidence is somewhat limited, there is currently no "gold standard" measure available for assessing knee-related function. The development of the instrument was supported by the literature and expert panel review; however, some of the factors

that were deemed most important to patients (e.g., satisfaction and activity limitations) were not included in the final instrument.

Furthermore, although this instrument has adequate content validity, there have been no studies published investigating the measurement properties at the item level. Some items, for example “Can you fully straighten your knee” and “Can you fully bend your knee” are relatively similar and may be redundant. While both of these items may be important for the clinician, the distinction between the two may be minimal for the patient. This is not an isolated incident with multiple examples of redundancy with questions. A detailed item analysis is necessary to ensure that all items are performing appropriately.

Another common concern related to test development is that an instrument is capable of measuring function across a wide spectrum of ability levels. The KOOS was originally developed and tested in the general population (predominantly involving ACL injuries). Because one of the primary purposes of this instrument is to measure long-term recovery following knee-related injuries (Roos et al., 1998) the instrument needs to be able to measure knee-related function at low levels (i.e., after an injury has occurred) as well as function at high levels (i.e., when a patient is prepared to resume normal daily activity). Although acceptable measurement properties have been demonstrated in the general population, it is unknown whether this instrument will still function properly in athletic populations where baseline physical function is greater (Muaidi et al., 2009).

Marx Activity Rating Scale

Activity levels of patients are important to consider because very active patients have different expectations and demands compared with relatively sedentary patients (Barber, Noyes, Mangine, & Hartman, 1990; Odensten, Hamberg, Nordin, Lysholm, & Gillquist, 1985). Because there is large variation in activity levels of patients, it is important to have a method for evaluating the activity levels of patients. Furthermore, activity limitations are a key dimension of the WHO ICF evaluation model.

Marx et al., conducted a systematic review to examine existing activity-rating scales, where five scales were identified (Marx, Stump, et al., 2001). Three of the five scales (Straub & Hunter, 1988; Tegner & Lysholm, 1985) classified activity based upon difficulty of the activities; however, difficulty was determined arbitrarily with no supporting rationale. The remaining two scales (Daniel et al., 1994; F. R. Noyes & Barber-Westin, 1997) classified activity by sport and activity demands. For example, Daniel et al., defined three levels of sport to classify patients (Daniel et al., 1994). Level one sports were described as jumping, pivoting, and hard cutting sports. Level two sports were considered sports that involve lateral motion, but with less jumping or hard cutting than level one. This can be problematic for patients that may participate in sports that require jumping, or pivoting, or both, but no cutting (e.g., track & field). None of the five scales included measures of frequency as well as intensity in their classification. Because of these limitations Marx et al., developed the Marx Activity Rating Scale (MARS) (Marx, Stump, et al., 2001). The purpose of the MARS was to

assess patients' activity levels for the purpose of discrimination across multiple time points. This instrument was not intended for use as an isolated measure, rather, as an additional measure in tandem with other outcome measures.

Development. Nine items were developed for the MARS through interviews with 15 content experts (10 orthopedic surgeons, 5 physical therapists and athletic trainers) and 20 athletic patients with knee-related injuries. These nine items included: getting out of a low chair, going up stairs, going down stairs, running, cutting, pivoting, jumping, decelerating, and doing a deep knee bend or a squat.

Fifty patients with existing knee injuries were asked to review the item list and rank by importance and severity. Average ratings by importance and severity were summed to create a total average score for each item. The top four items were selected based upon total average score: running, cutting, decelerating, and pivoting. The authors reported that four items were chosen because that is the length they wanted. No additional rationale was provided for why the items were added.

An additional review, including 25 athletic patients with existing knee injuries, occurred to revise items for clarity and understanding. Although the authors specified that modifications to instructions and items were made following each administration, it was not reported how many revisions were necessary to create the finalized version of the MARS.

Scoring System. All items were evaluated with ordinal scales using 5-category response options with frequency-related descriptors. Categorical

options were consistent for all items (0 = “Less than one time in a month”; 1 = “One time in a month”; 2 = “One time in a week”; 3 = “2 or 3 times in a week”; and 4 = “4 or more times in a week”). Scores for the MARS range from 0 to 16, where a higher score indicated a higher level of physical activity.

Measurement Properties. Reliability of the MARS was assessed using ICCs with a test-retest interval of one week in 40 athletic volunteers. Above-average ICC values were noted between test intervals (ICC = 0.97). The type of ICC model used for analysis was not specified.

Evidence of convergent validity was assessed for the MARS through correlations with existing activity scales (Tegner scale $r = 0.66$; Cincinnati scale $r = 0.67$; Daniel scale $r = 0.52$) (Marx, Stump, et al., 2001).

Responsiveness was not assessed for the MARS.

Relationship to ICF. The MARS can be used to assess limitations in activity. Because of the small number of items included in the instrument, this is the only domain that is assessed. This supports the decision of the authors to use this instrument as an add-on to existing knee-related PRO, rather than as a stand-alone instrument.

Additional Limitations. The MARS provides a self-reported estimate of a patients’ perceived activity level. This information could be useful in helping to assist clinicians with injury-related decision-making. The score system however, is severely limited. Some activities, such as equestrian show jumping (competitive sport of riding horses over a course of fences and other obstacles in an arena, with penalty points for errors) result in high physiological demands,

specially related to the trunk and lower extremities, yet do not require jumping, running, pivoting, or cutting related activities (Douglas, Price, & Peters, 2012). A patient whose predominant form of activity is equestrian show jumping would be identified as having significantly low or even zero activity despite the high physical demands of the sport.

International Knee Documentation Committee Subjective Knee Form

The IKDC Standard Evaluation Form, though widely used, primarily emphasized clinical-based outcomes. Because of evidence indicating the benefits of PRO, the IKDC in collaboration with the American Orthopaedic Society for Sports Medicine developed the IKDC-SKF. The purpose of this instrument is to detect changes in symptoms, function, and sports activity experienced by patients with a variety of knee conditions (Irrgang et al., 2001).

Development. The IKDC proposed an initial set of 48 questions related to symptoms ($n = 27$), function during activities of daily living ($n = 8$), function during sports activities ($n = 4$), current function of the knee ($n = 3$), participation in sports, work activities, or both ($n = 5$) and mood ($n = 1$). A general overview of the content within each domain is displayed in table 4.

Table 4. International Knee Documentation Committee Subjective Knee Form
Question Domains and Content

Domain	Content
Symptoms	• Pain • Swelling • Stiffness • Giving way • Locking • Walk on level surfaces • Ascend and descend stairs • Stand • Kneel on the front of the knee • Squat
Function during activities of daily living	• Sit with the knee bent • Rise from a chair • Run straight ahead • Jump and land on the involved leg
Function during sports activities	• Stop and start quickly • Cut and Pivot

The original 48-item instrument was pilot tested in a sample of 144 patients currently receiving treatment at orthopedic sports medicine practices. Pilot-test data were used to revise\delete existing items and to develop new items. Beyond stating that many of the items were frequently unanswered (e.g., pain, swelling and presence of giving way), the authors did not provide any descriptions of the original items or explain which items were revised/deleted.

The revised item set contained 41 items related to symptoms ($n = 19$), function during activities of daily living ($n = 8$), function during sports activities ($n = 4$), current function of the knee ($n = 3$), participation in sports or work activities ($n = 5$), mood ($n = 1$), and overall health ($n = 1$). The revised item set was evaluated in a sample of 222 patients currently receiving treatment at orthopaedic sports medicine practices (same centers as the previous revision). A Rasch partial credit model analysis was used to evaluate measurement properties of the items and eliminate items with poor fit indices (author selected criteria: infit statistics < 0.6 or > 1.4). After removal of poor fitting items, the final version of the IKDC-SKF included 19 items related to symptoms ($n = 7$), sports activities ($n = 10$), and function ($n = 2$) (Irrgang et al., 2001).

Scoring System. To determine the best method for scoring the subjective evaluation form three separate scoring methods were developed and evaluated. The three different methods included: [1] adding the scores for all items together for a single total score; [2] adding the weighted scores for items from each domain; and [3] scoring based upon item response theory. To compare the efficiency of each scoring method Pearson correlation coefficients were calculated to compare the weighted scoring method and the item response theory (IRT) scoring method to the simple summation method.

The correlation coefficient between the weighted scoring system and the summation method was 0.99. The correlation coefficient between the IRT scoring method and the summation method was 0.98. In addition to the high correlations between scoring methods, no differences were found in the five highest and five

lowest scoring patients regardless of scoring method. As a result, the summation scoring method was selected due to its simplicity and high correlation with the more complex scoring methods. On the IKDC-SKF the maximum sum score is 87. The response to item 10a “Function Prior to Knee Injury” is not included in the overall score. Scores for the IKDC-SKF are then re-scaled to a 100-point scale; where higher scores represent lower levels of symptoms and higher levels of knee-related function (Irrgang et al., 2001).

Measurement Properties. Evidence of convergent validity was assessed through correlations with summary scores from the SF-36 generic-function instrument. The IKDC-SKF scores were poorly correlated with SF-36 mental-summary scores ($r = 0.16$) and weakly-moderately correlated with SF-36 physical-summary scores ($r = 0.66$).

Reliability for the IKDC-SKF scores were assessed in two studies using intraclass correlation coefficients. Coefficients in the two studies ranged from 0.82 – 0.95 (Irrgang et al., 2001; Paxton et al., 2003).

Responsiveness for the IKDC-SKF was assessed with follow-up examinations at 6-months and 12-months.

Relationship to ICF. Body functions and structures is assessed through the seven symptom-related questions. Activity limitations are assessed through nine of the sports activities questions. Similar with the KOOS however, many of these items represent low levels of function (e.g., “go upstairs”, “sit with your knee bent”, “rise from a chair”). No items are included to assess environmental factors.

Additional Limitations. Although this instrument was developed using sound measurement principles, there are some concerns related to the scoring system. In the previous development of the IKDC standard evaluation form the authors explained that a major limitation of the clinical outcome measures was the lack of rationale for scoring. Patients with limitations in one section could still obtain a relatively high score due to higher scores in other categories (E. Hefti et al., 1993). This theory was discarded in the revised PRO instrument where scores are based on a summed total from all categories. While various scoring methods were assessed, each of these scoring methods used total score from all domains.

Limitations of Patient-Reported Outcomes Instruments

Interpretation of scores from an instrument are highly dependent upon the validity and reliability of the instrument. When either of these measurement properties are low the scores may not be representative of the traits (i.e., ability) they were designed to measure. Ability in the context of knee-related PRO instruments relates to knee-related function. Development of an instrument should be based in theory with evidence to support the content and population specific validity of an instrument as well as reliability. Because the primary purposes of PRO instruments are to discriminate patients at various levels of function and assess changes in function over time, responsiveness is also an important measurement consideration. The following section describes some of the limitations of the current instruments, specifically related to content validity, population validity, and ceiling effects.

Content Validity

Function as described by the WHO is a multidimensional construct which has been described using the ICF model framework. In order to ensure that instruments are accurately describing function, items assessing each domain of the ICF are necessary. Despite their use in clinical practice, none of the recommended knee-related PRO instruments are capable of assessing all domains of function as described by the WHO ICF model framework.

Part of the reason for the lack of sufficient item coverage is that many of the recommended instruments were developed before the proposed ICF model was released. Early model frameworks such as the Nagi and ICIDH did not include environmental factors as a key component of measuring function and disablement. Revisions to the current instruments are necessary to develop these items. Functional limitations are not simply a feature of the individual, rather they are an outcome of an interaction of the person with a health condition and environmental factors (Schneidert, Hurst, Miller, & Üstün, 2003).

The role of the environment can be challenging to assess because it influences both environmental barriers and facilitators that have an effect on a person's level of function. Environmental barriers relate to the challenges a person experiences related to engagement in life situations. For example, a patient who lives on the second floor of an apartment complex is likely to perceive activity limitations related to ascending and descending stairs differently than someone who lives on the first floor, or someone who has access to elevators where stairs may represent a minor inconvenience. Living with a

significant other, who is capable of providing assistance with daily activities, would be considered a form of environmental facilitation. Because of the vast array of environmental barriers and facilitators it can be extremely challenging to measure this construct; none the less, it is a vital component of function. This information can help to guide treatment options for optimizing treatment for patients and their families.

Previous studies have attempted to examine environmental influences on function using the movement and activity in physical space (MAPS) system (Farnsworth II, McElhiney, David, Sinha, & Ragan, 2013; Herrmann et al., 2011). The MAPS score, a component of the MAPS system, is derived from accelerometer and GPS data to assess limitations in activity, participation, and the environment. While the results of the MAPS score have been successful, it requires substantial burden on the clinician to analyze and interpret the data. In the clinical setting more time efficient methods for evaluating environmental influences are necessary. The SF-36 includes a number of items related to role-functioning (i.e., environmental factors). Inclusion of these items as part of current knee-related PRO may provide an opportunity to assess environmental factors with minimal clinician-patient burden.

Population-Specific Validity

Scores from an instrument are considered valid when they accurately represent the trait they were developed to assess. Population-specific instruments are designed to be used with specific populations; however, to some extent, all instruments are designed with some population in mind. Even

instruments designed to be used for the general population, may not be appropriate for all specific sub-populations.

For example, the MARS was developed to assess physical activity levels in adults. When evaluating physical activity in patients 18 years and older the instrument has high reliability measured by ICC. A recent study, however, identified that adolescents under the age of fourteen had poor test-retest reliability (Shirazi, Israel, & Kaar, 2016). Measurement properties of an instrument are population specific. For this reason, it is necessary to evaluate the properties of an instrument in the population they will be used with. In a systematic review of the KOOS, evidence suggests that some of the scales of the KOOS may not be relevant or useful in all populations. In older adults the activities of daily living subscale had better content validity, while the sports/recreation subscale was more appropriate with younger patients (N. Collins et al., 2016).

In sports medicine PRO instruments are necessary to document changes in function as a patient recover from injury. The KOOS and IKDC-SKF are commonly used to assess knee-related function. To ensure their appropriateness for use in athletic populations, these instruments must be capable of assessing a wide range of function. When an athlete is injured, their function levels will be low, similar to the general population. Athletes, however, are presumed to have a higher functional threshold compared to non-athletic individuals (Muaidi et al., 2009).

Research in motor learning has demonstrated that although individuals may possess different level of innate ability, training and repetition will improve motor skills (e.g., dynamic strength, speed of limb movement, multi-limb coordination, control precision) (Coker, 2013). This suggests that athletes, who are continuously engaged in demanding physical activities, are likely to have higher levels of function compared to non-athletic counterparts. If PRO instruments, which were developed for use in non-athletic and general populations are used with athletes it is possible that the instruments may not be able to effectively evaluate knee-related function, particularly when function is high. Relatively few studies have investigated the measurement properties of these instruments in athletic populations (Marx, Jones, et al., 2001).

Ceiling Effects

Ceiling effects are a known limitation of many PRO instruments (Fries et al., 2011; Roos & Toksvig-Larsen, 2003). Significant ceiling effects are a problem because it means the instrument is not able to fully measure a patient's function, especially the upper ranges of ability. Patient reported outcomes, which are often used to assess changes in function related to treatment, become less useful as a patient reaches the upper limit of the tests (a.k.a. ceiling effect). An example of this phenomenon is illustrated in figure 5.

In this example, a patient's knee-related function is evaluated multiple times following an injury to assess treatment efficiency and determine when it is safe to resume normal activity. As expected, the patient's knee-related function score is low after injury (observed score 1) and improves over time as the injury

heals (observed score 2 & observed score 3). After the third assessment; the instrument is no longer able to measure changes in function because there are not items capable of measuring the patient's function. In this scenario, the patient may be returned to normal activity too soon, before their function has returned to pre-injury levels.

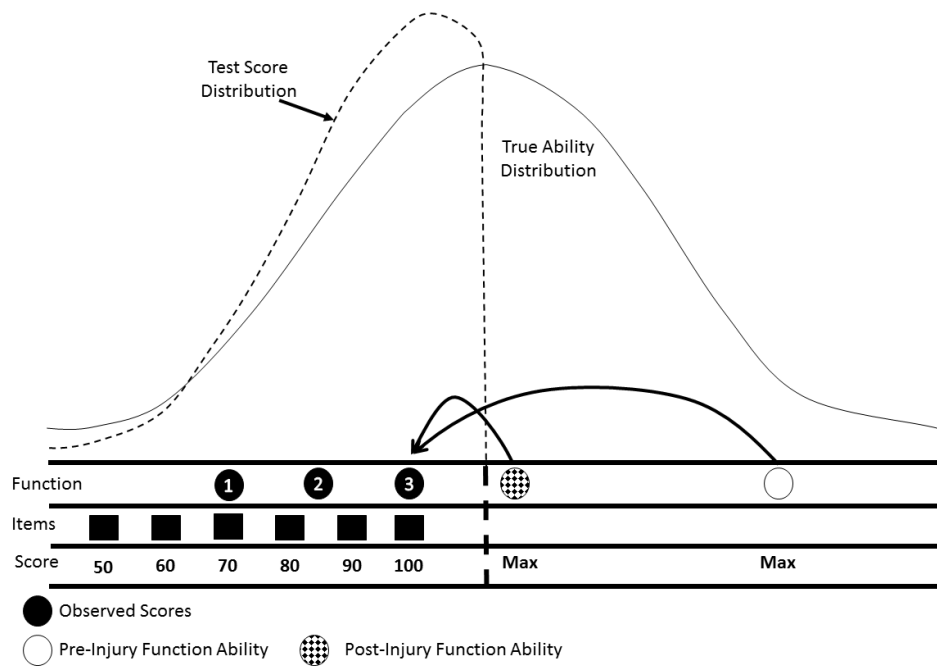


Figure 5. Theoretical Knee-Related Function Distribution. Adapted from Ragan, BG (2004). An empirical investigation of several critical psychometric issues in neuropsychological testing of mild traumatic brain injuries. University of Illinois at Urbana-Champaign, Urbana, IL.

Prevalence of Ceiling Effects

Twelve months following total knee replacement ceiling effects were identified in the pain (15 – 22%), sport/recreation (16%), and quality of life (17%) sub-scales of the KOOS (Roos & Toksvig-Larsen, 2003). In addition, a 10% ceiling effect was noted with the KOOS in adults following treatment of focal cartilage lesions. Furthermore, preliminary pilot data collected on 79 adults with existing knee injuries revealed ceiling effects ranging from 3 – 28% across subscales of the KOOS. These ceiling effects were identified in normal adult populations. In athletes where function is presumably higher, it is likely that these problems would be more pronounced.

Marx et al., conducted an evaluation of the measurement properties of various PRO measures (the Activities of Daily Living Scale, the American Academy of Orthopaedic Surgeons sports knee-rating scale, the Cincinnati knee-rating scale, the Lysholm scale, the Tegner scale, and the Short-Form 36) in injured athletes. In their review Marx et al., claimed there were no ceiling effects in any athletes prior to their injury (Marx, Jones, et al., 2001). Despite these claims, many of the instruments had reported “highest scores” equivalent to the maximum score of each respective instrument. In addition, potential floor effects were identified in a few of the measures where the “lowest score” was 0. Z-score transformation of baseline function scores across outcome measures suggest that a range of 1.22% – 9.48% of athletes obtained the maximum score.

Performance-Based Outcomes: A Measure of Higher-Level Function?

Following injury, knee-related function will be low. As patient function improves, many clinicians begin using performance-based outcomes measures to better simulate real world functional ability. Examples of performance-based measures include timed chair stands, single-leg stand, grip strength, elbow flexion and knee extension strength (strength chair), and the six-minute walk (Curb et al., 2006; Terwee, Mokkink, Steultjens, & Dekker, 2006).

A systematic review of performance-based outcomes related to osteoarthritis of the knee and hip indicated that many performance-based outcomes failed to demonstrate acceptable measurement properties for reliability, validity, and responsiveness (Terwee et al., 2006). Furthermore, practical limitations such as insufficient time, limited space, and inadequate equipment limit the use of these outcomes in every day clinical practice (Terwee et al., 2006). In addition, these measures are limited in their ability to evaluate the domains of function as specified in the WHO ICF model.

Improving Patient-Reported Outcomes

Establishment of best treatment practices require documentation of outcomes throughout the recovery process. Limitations of the current instruments limit their ability to evaluate function, as defined by the WHO, throughout the recovery process, specifically relating to insufficient evaluation of environmental factors, lack of sufficient evidence supporting use in athletic populations, and ceiling effects. These issues must be addressed to improve the current

instruments and make them more clinically relevant. Development of new items that [1] evaluate environmental factors related to function, [2] minimize ceiling effects, and [3] function properly in athletic populations are desperately needed.

Test Construction & Item Development

Varieties of psychometric methods are available to support the development and evaluation of PRO instruments. The most common methods include Item Response Theory (IRT), Rasch Measurement Theory, and Classical Test Theory (CTT). Each of these various methods have advantages and disadvantages. The method best suited for test development and evaluation is dependent upon the intended uses of the instrument. The following section briefly outlines these methods as well as their advantages and disadvantages for use with this study.

Item Response Theory

IRT is an advanced statistical procedure that can be used to improve knee-related PRO. IRT has been used previously to evaluate and improve PRO such as the PRO measurement information system (PROMIS) function instruments. In this section, the fundamentals of IRT, models, historical overview, characteristics and assumptions of IRT, information functions, relative efficiency, and advantages over classical test theory are addressed.

Fundamentals of Item Response Theory

IRT is used to evaluate the characteristics of items and the ability of a person to determine the likelihood of responding in a particular way to an item from an instrument. With IRT, we are interested in the probability of a patient to

respond correctly to an item, in the case of dichotomous items; or responding to a particular category, in the case of polytomous items. There have been many different models developed to evaluate dichotomous and polytomous items.

The basic premise of IRT is the investigation of the relationship between a patient and their response to an item. The probability of answering an item correctly is dependent upon the item difficulty (b) and the patient's ability level (θ). A patient has a 50% likelihood of responding correctly to an item when ability matches the difficulty of the item. When patient ability is high relative to the item's difficulty the likelihood of a correct response increases. Conversely, when ability is low, relative to item difficulty, the likelihood of a correct response decreases. The relationship between item difficulty and patient ability can be represented mathematically using the function:

$$P(\theta) = \frac{e^{(\theta-b)}}{1 + e^{(\theta-b)}} \quad (1)$$

where $P(\theta)$ is the probability that a patient with ability θ responds correctly to an item of difficulty b . The values of θ and b can range from negative infinity to positive infinity and e is the natural logarithmic base ($e = 2.718$).

The response probability is converted to a log-odds scale similar in concept to a z-score scale. The log-odds can be represented with the following expression:

$$O_i = \frac{P_i(\theta)}{Q_i(\theta)} \quad (2)$$

where O_i is the odds of a correct response and $Q_i(\theta)$ is equivalent to $1 - P_i(\theta)$ or the probability of an incorrect response. In other words, the log-odds is the probability of a correct response over the probability of an incorrect response. In its most basic form, the probability of an incorrect response on item i is represented by the expression:

$$Q_i(\theta) = \frac{1}{1 + e^{(\theta - b)}} \quad (3)$$

As a results, the odds of a correct response can be represented by the expression:

$$O_i = e^{(\theta - b)} \quad (4)$$

The natural logarithm of equation (4) gives us the expression:

$$O_i = \ln(\theta - b) \quad (5)$$

Thus, the odds of a correct response for item i is the natural log of the difference in ability and item difficulty. The units on the log-odds scale for IRT is the logit, which is a contraction of the words *log odds unit*. In essence, the log-odds scale, which has ratio-scale measurement properties, represents the probability that a person of ability level θ for a given trait (e.g., function) will answer an item of difficulty level b correctly. This is the fundamental principle of the one-parameter IRT and Rasch simple logistic models (Lord, 1980).

The logit function follows a linear model where logits are equal distances apart with ability and item difficulty represented on the same line. Visually the logit function follows an asymptotic distribution (S-shaped curve) also known as the item characteristic curve (see figure 6) (Lord, 1977).

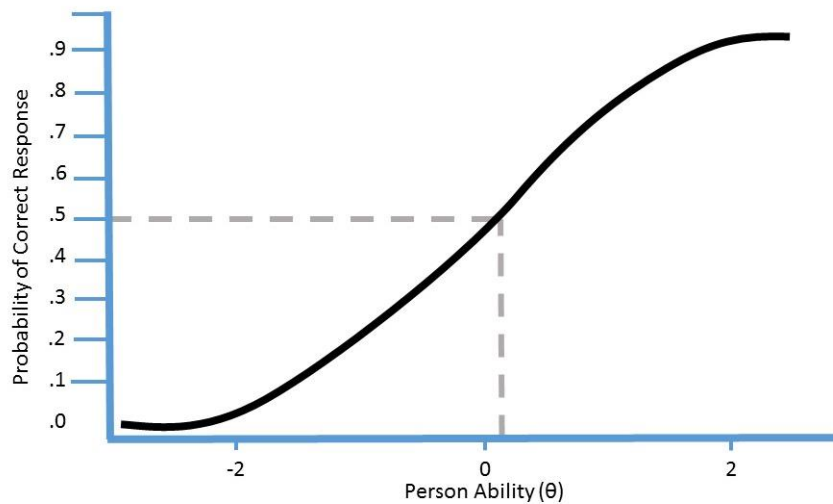


Figure 6. Item Response Theory Item Characteristic Curve. The x-axis indicates person ability (θ) on a continuum from negative to positive infinity. The y-axis indicates the probability of a correct response ranging from 0 – 1. The location of the curve is determined by the difficulty of the item where a person with ability θ has a 50% probability of a correct response.

One-parameter Item Response Theory and Rasch simple logistic

models. The one-parameter IRT model and the Rasch simple logistic model are mathematically equivalent and assume that all items have equal discriminating power and there is minimal chance of guessing an item correctly (Hambleton, 1991; Lord, 1980). Although mathematically similar these two models were developed independently and follow distinctly different philosophical differences. The basic formula for the one-parameter IRT and Rasch simple logistic models are both mathematically similar to equation (1) described previously. The one-parameter model is expressed as follows:

$$P_{ij}(\theta) = \frac{1}{1 + e^{-(\theta_i - b_j)}} \quad (6)$$

where $P_{ij}(\theta)$ represents the probability of a correct response for person i on item j ; θ_i represents the ability of person i ; b_j represents the difficulty of item j , for the one-parameter IRT and Rasch models (equation 6) (Lord, Novick, & Birnbaum, 1968; Rasch, 1961).

The item difficulty parameter is very important in the visualization of the logit function and influences the location of the item characteristic curve. A difficult item, compared to a less challenging item will require a higher ability to achieve the same probability of a correct response. Thus, the item characteristic curve for the more challenging item will be shifted to the right of the easier item. As the difficulty of an item increases a higher ability is required to achieve the same probability of guessing the item correctly, compared to an easier item. This

is illustrated in figure 7, where the item characteristic curve for item k , the most challenging item, is seen further to the right of items j and i , due to their lower difficulty.

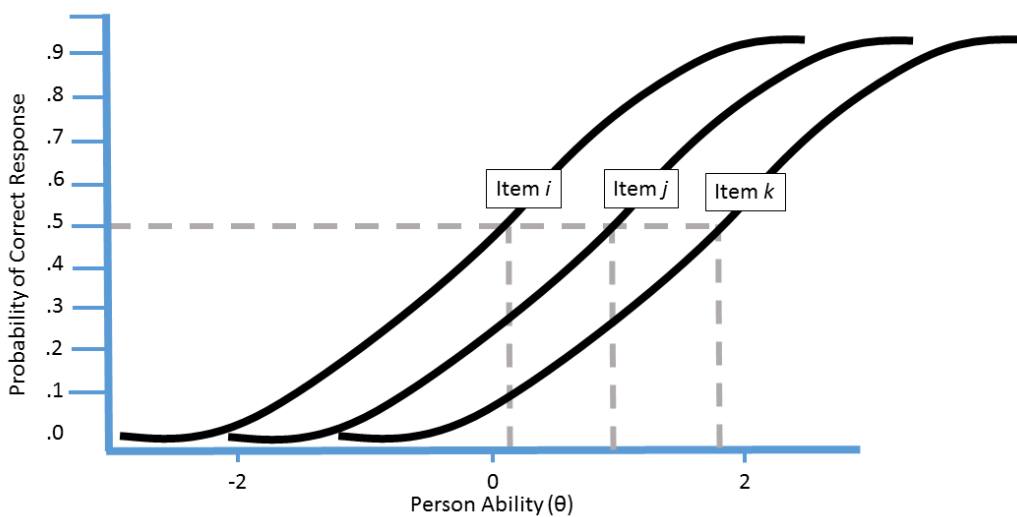


Figure 7. Item Characteristic Curve (1-parameter model). Item characteristic curves for three items i , j , & k . The location of the curve for each item is dependent upon the difficulty of the item where a person with ability θ has a 50% probability of a correct response. Items of lower difficulty will be shifted left, relative to a more challenging item.

Two-parameter Item Response Theory model. The two-parameter IRT model is different than the one-parameter IRT model because in this model, all items are not assumed to have the same discriminating power. For the two-parameter IRT model an additional parameter is added the item discrimination parameter (α). Item difficulty (b) is measured and defined as it was in the one-parameter IRT model. The item discrimination parameter is related to the slope of the item characteristic curve. Mathematically the two-parameter model is expressed by the function:

$$P_{ij}(\theta) = \frac{1}{1 + e^{\alpha_j(\theta_i - b_j)}} \quad (7)$$

where $P_{ij}(\theta)$ is the probability of a correct response for person i on item j ; α_j is the item discrimination parameter and $(\theta_i - b_j)$ is the difference in ability of person i and the difficulty of item j . The α parameter is represented graphically by the slope of the item characteristic curve. Figure 8 illustrates two items (i & j) with the same difficulty parameter, but different item discrimination parameters. The slope for item i is very steep, which indicates higher discriminating ability compared to item j . An item with a high discriminating power is better able to distinguish between high and low ability persons for a specific level of θ .

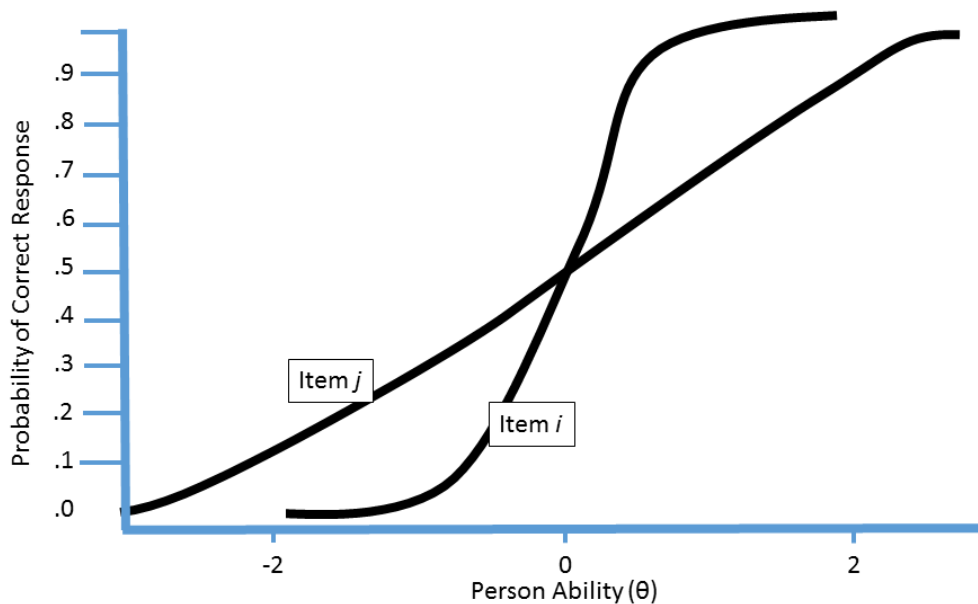


Figure 8. Item Characteristic Curve (2-parameter model). Item characteristic curves for two items with equivalent item difficulty parameters (b), but different item discrimination parameters (α). The item discrimination parameter for item i is greater than the item discrimination parameter for item j . This can be seen in the slope of the two item characteristic curves, where the slope for item i is greater than the slope for item j

Three-parameter Item Response Theory model. The three-parameter IRT model is the most complex of the three models. In addition to the item difficulty and item discrimination parameters, the three-parameter IRT model includes an additional pseudo-guessing parameter. This model is most appropriate for use when there is a chance that examinees can answer an item

correctly by guessing, such as the case with multiple-choice questions. The pseudo-guessing parameter, which is represented by the lower asymptote, indicates the likelihood of a patient responding correctly to an item due to random chance. Mathematically the three-parameter IRT model can be expressed by the function:

$$P_{ij}(\theta) = c_j + \frac{1-c_j}{1+e^{-\alpha_j(\theta_i-b_j)}} \quad (8)$$

where $P_{ij}(\theta)$ is the probability of a correct response for person i on item j ; c_j is the pseudo-guessing parameter; α_j is the item discrimination parameter and $(\theta_i - b_j)$ is the difference in ability of person i and the difficulty of item j . Figure 9 illustrates two items with different pseudo-guessing parameters. As seen in the figure, item i has a greater pseudo-guessing parameter compared to item j , which is reflected in the lower asymptotes of the two items and their position relative to one another. The pseudo-guessing parameter is important for items where there is a possibility of guessing an item correctly through random chance alone. The random error may inflate the number of correct responses leading to inaccurate parameter estimates. The additional item parameter attempts to control for this random error.

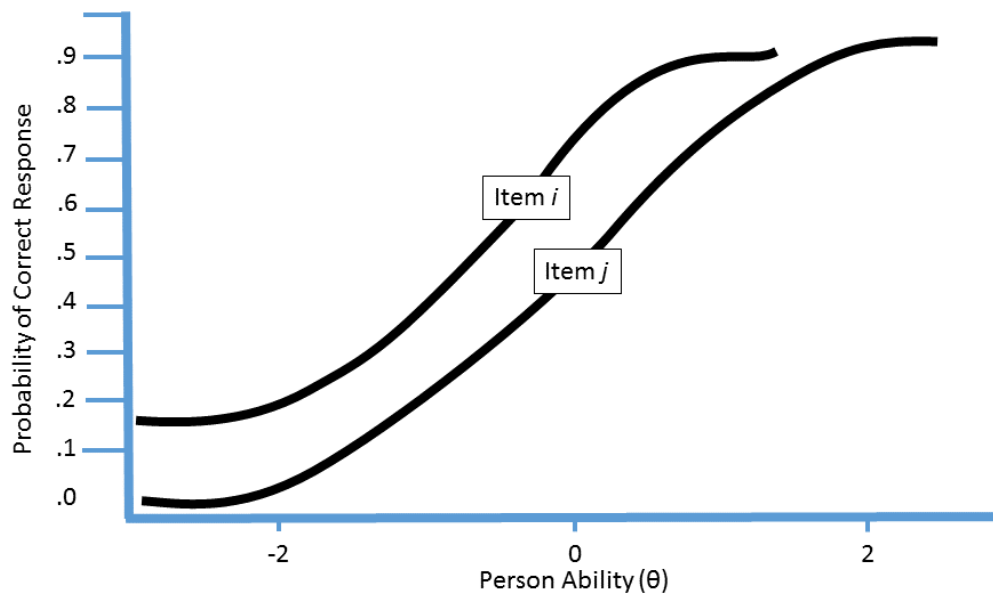


Figure 9. Item Characteristic Curve (3-parameter model) Item characteristic curves for two items with different pseudo-guessing parameters. Item *i* has a higher pseudo-guessing parameter compared with item *j*. This can be seen due to the difference in lower asymptotes between the two items. The greater the lower asymptote the higher the probability of a correct response due to random chance.

Item Response Theory Model Characteristics and Assumptions

Item Response Theory is a collection of advanced statistical models used to evaluate the probability of success on a given item. The following section describes some of the fundamental characteristics of IRT and basic assumptions.

The assumptions of IRT include dimensionality of the test, local item independence, and parameter invariance.

Item Characteristic Curve. Item characteristic curves are the backbone of IRT. The item characteristic curve is a non-linear (logistic) regression line, where the item performance is regressed across patient ability (Lord, 1977). Thus, for a given item, the item characteristic curve indicates the probability of success for each level of theta across the continuum of ability scores. Logistic functions model the probability of success and follow a monotonically increasing function, such that higher ability results in a higher probability of success. For example, an athletic trainer is more likely to respond correctly to a question related to injury prevention compared with a parent or layperson.

Dimensionality of the instrument. One of the major assumptions of IRT is that all items in the test are only measuring a single latent trait. Unidimensionality is necessary because of the mathematical models used for item analysis. When multiple dimensions are included in a single instrument, item-parameter and person-ability estimates are likely to be biased.

In practice, unidimensionality of an instrument may not be feasible. Experts have argued that responses to items are multiply determined (Hambleton, Swaminathan, & Rogers, 1991; Reckase, 1979) (e.g. caused by more than one factor), meaning several minor abilities are required to respond to items. For example, an instrument designed to assess knee-function may ask an individual about their ability to ascend and descend stairs. Although knee-related

function is the predominant trait being assessed, ankle-related function may play a small role in their ability to ascend and descend stairs.

Elimination of secondary abilities may not be possible during the development of instruments. Therefore, it is important to assess the dimensionality of an instrument to determine which ability is the “dominant” factor or trait.

The dimensionality of an instrument can be assessed through a number of different methods; however, the use of Cronbach’s alpha is perhaps the most widely used method (Hattie, 1985). According to Cronbach a high alpha coefficient, which is the sum of all split-half coefficients, indicates a “high first factor saturation”, or that alpha was an index of “common factor concentration” (Cronbach, 1951). Alternative approaches to evaluating dimensionality include the DIMTEST (Stout, 1987), Holland and Rosenbaum’s approach (P. R. Rosenbaum, 1984), factor analysis (Conway & Huffcutt, 2003), and principle components analysis of residuals, which have all been shown to correctly identify unidimensionality in simulation studies (Nandakumar, 1994).

Violations of unidimensionality can have significant influences on item-parameter and person-ability estimates. In simulation studies it was found that when multiple highly correlated dimensions were present, ability estimates were correlated with the factor scores for each dimension (Reckase, 1979). When dimensions are equally prominent, ability estimates may represent the sum or average of ability estimates for each dimension. In addition, only slight deviations occurred between *true* and estimated item-parameters in multidimensional models (Dorans & Kingston, 1985; Drasgow & Parsons, 1983). As the

correlations between dimensions decrease, item-parameter estimates grow more unstable (Ansley & Forsyth, 1985).

Local dependence. Local independence is another important assumption for IRT (Lord, 1980). The assumption of local independence dictates that a person's response for a given item is dependent on their level of ability related to the measured trait. Violations can be either trait dependent or response dependent. Local independence is important because it can influence the interpretation of item-parameter and person-ability estimates. When local independence is true, the probability of answering a pair of items correctly can be determined from the product of the probabilities of each item. This is expressed mathematically in equation 9.

$$P_{ij}(\theta_k) = P_i(\theta_k)P_j(\theta_k) \quad (9)$$

Trait dependence is associated with the dimensionality of the tests. When items are multi-dimensional, meaning multiple traits are being assessed, the probability of a correct response is influenced by both abilities. Relating to local independence, this means that something other than the trait being measured is influencing response patterns; thus, the local independence assumption is violated. Many instruments are instructed which intentionally violate trait dependence. For example, many of the tests designed to assess knee-related function include multiple subscales (e.g., pain, activities of daily living, sport & recreation). Computerized neurocognitive tests for assessing cognitive function are another example, which often include multiple domains of cognitive function

(e.g., verbal memory, reaction time, working memory). If the two traits are highly correlated then this violation will have minimal effects on a patient's response patterns (Dorans & Kingston, 1985; Drasgow & Parsons, 1983).

Response dependence is another type of violation. Response dependence occurs when another item influences a patient's response to an item. An example of response dependence is found when the response of one item provides a clue or hint about the answer to a later question. While trait dependence is often deliberate, response dependence is more likely the result of inappropriate item construction. An example of response dependence is the physical functioning subscale of the SF-36, where the items *Climbing one flight of stairs* and *Climbing several flights of stairs* are dependent on one another (Kreiner & Christensen, 2007). Another example in the same instrument includes; *Walking one block*, *Walking several blocks* and *Walking more than a mile* (Kreiner & Christensen, 2007).

Violations of local independence are important because they can inflate reliability and result in inaccurate estimation of item parameters, test statistics, and person ability (Fennessy, 1995; Sireci, Thissen, & Wainer, 1991; Thissen, Steinberg, & Mooney, 1989). Additionally, local item dependence may introduce an additional dimension into the test at the expense of the construct of interest (Wainer & Thissen, 1996). A potential solution for addressing violations of local independence is to combine dependent items into a single polytomous item (Marais & Andrich, 2008). This approach is successful at providing more

accurate presentation of parameter estimates when violations of local independence are present (Marais & Andrich, 2008).

Parameter Invariance. The invariance of person's and items is a major advantage of IRT models and has direct implications for computer-adaptive testing, cognitive diagnostic assessment, and test equating (Embretson & Reise, 2013; Hambleton et al., 1991). In IRT, item and ability parameters are said to be invariant. The property of invariance is obtained by incorporating information about the items into the ability-estimation process and by incorporating information about the patients' abilities into the item-parameter-estimation process (Hambleton et al., 1991). This fundamental function of IRT is important because it means that ability estimates obtained from different sets of items will be the same, and item parameter estimates obtained from different groups of examinees will be the same. The invariance of item parameters has been illustrated in figure 10, which shows the ability distributions for two groups of examinees. The examinees have the same probability of giving a correct response to the item, regardless of whether they are from group 1 or group 2. Because the probability of success is the same for both groups, the item-parameters must be the same for each group.

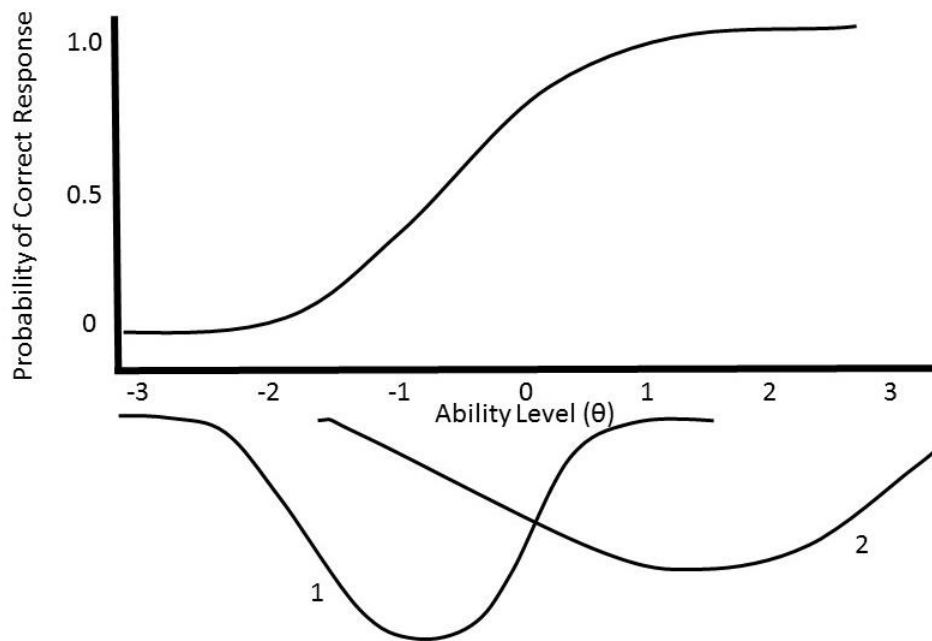


Figure 10. An item characteristic curve and ability distribution for two groups of examinees. Adapted from Hambleton RK. *Fundamentals of item response theory*. Vol 2: Sage publications; 1991.

When measuring function, invariance of item-parameters and person-ability estimates is especially important. As athletes undergo treatment following injury, their function (i.e., ability) is expected to change. Thus, the invariance property of IRT allows the same instrument to be used to measure a patient throughout the entire recovery process, if sufficient items are available to measure each level of function.

Item and Test Information Functions

In addition to invariance, IRT provides estimates of standard errors for individual ability estimates, rather than a single estimate of error for all patients. These conditional standard errors describe how much error is associated with the probability of success for a given ability level. The amount of information obtained for a given item can be determined mathematically from:

$$I(\theta) = p_i(\theta)q_i(\theta) \quad (10)$$

where $p_i(\theta)$ is the probability of a correct response for a given ability level; $q_i(\theta)$ is the probability of an incorrect response; and I is the information function. It can be seen mathematically that the greatest amount of information for any particular item will be provided when the ability of the person is equivalent to the difficulty of the item; where probability of success will be .5 (i.e., $.5 \times .5 = .25$). As the probability of success increases or decreases the amount of information obtained from each item decreases as well. For example, if the probability of success were .7; then the probability of an incorrect response would be .3. The item information function for this case would be .21.

The amount of error associated with the probability of success for a given item is represented by the reciprocal of the item information function.

$$SE(\theta) = \frac{1}{\sqrt{I(\theta)}} \quad (11)$$

where $I(\theta)$ represents the item information function as expressed in equation 10. This equation is an extension of traditional reliability functions and can help determine the value of items for a given level of theta.

When the ability distribution for a particular population is known, the item information function can be a valuable asset to select the items that are most appropriate for evaluating the population. If information is only needed for a specific ability range; for example, determining the readiness of an athlete to return to play following a knee injury; items can be selected that are capable of assessing the appropriate ability range while minimizing measurement error.

Figure 11 below illustrates the item information function for four unique function-related items (e.g., items “a”, “b”, “c”, & “d”). Immediately following injury, item “a” would be the most useful because it is expected that function ability would be low. On the other hand, item “d” would be the most informative when determining readiness to return to play. Item “c” provides almost no meaningful information related to player ability and may not be necessary.

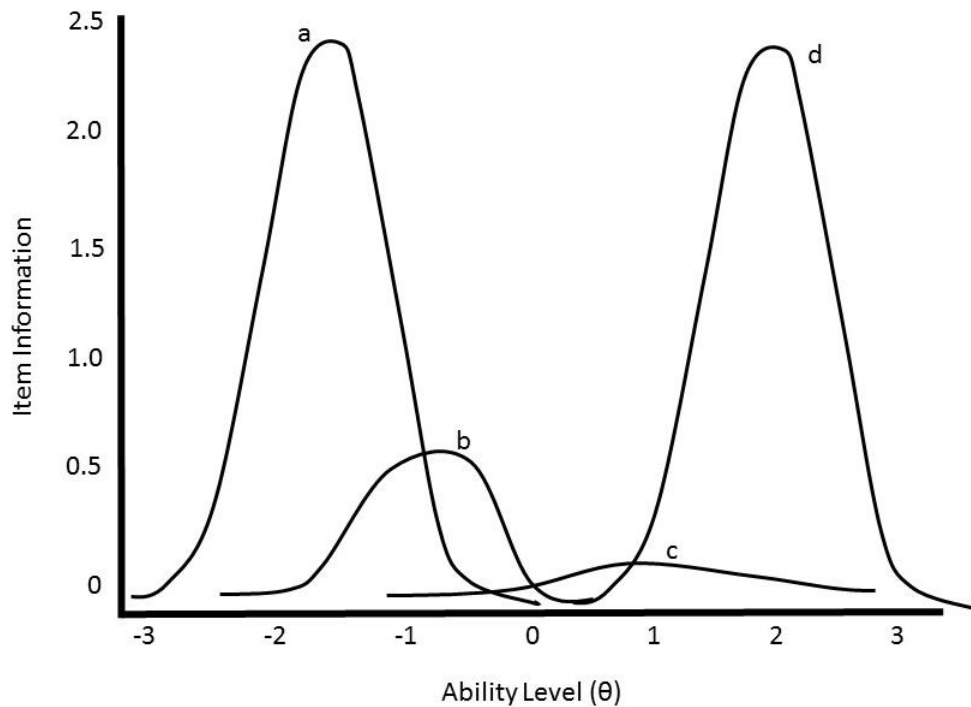


Figure 11. Item Information Function for Four Items (“a”, “b”, “c”, & “d”).

Alternatively, the item information function can also be used to determine if additional items are needed to improve assessment for a given ability range. In figure 12, we can see that there is relatively few information provided by these items for participants with $\theta \approx 0.5$. This gap in item coverage could result in poor estimates of ability for patients near this range of ability, where patients' ability may be over or under estimated. The total amount of information provided by the items of a test can be determined by the sum of the item information functions. The resultant distribution is the test information function. Figure 12 provides an

illustration of the test information for the four items in figure 11. As seen in the figure, there is a gap in the information provided at $\theta \approx 0.5$. Because of the invariance property of IRT, additional items can be added as needed to improve measurement of ability where gaps in the test information function are identified.

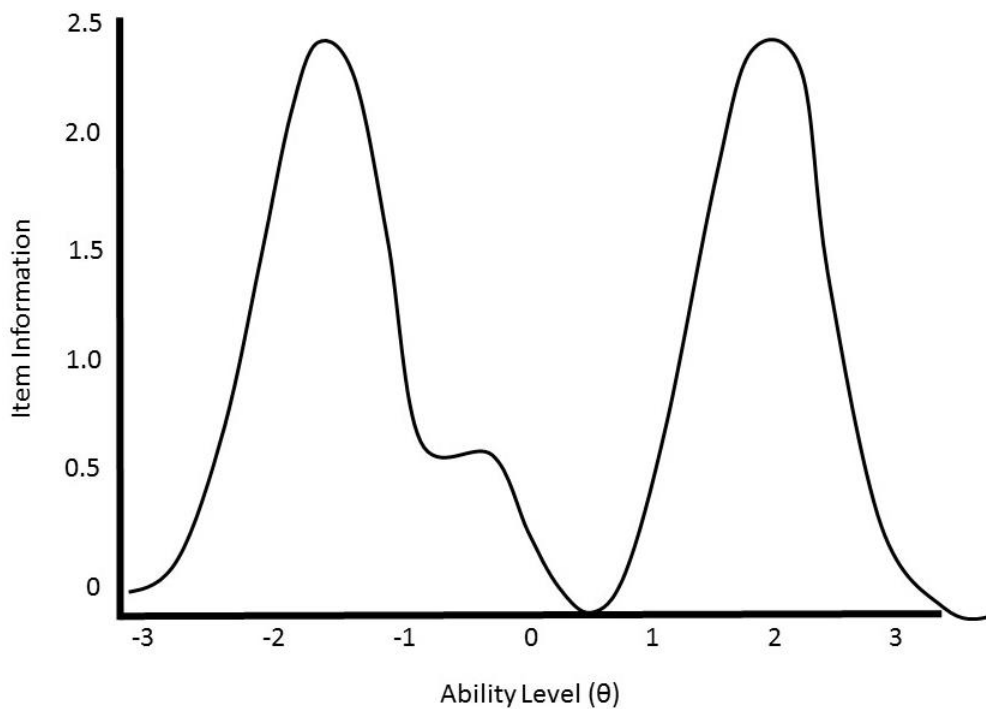


Figure 12. Test Information Function for Four Items (“a”, “b”, “c”, & “d”). The ability of the instrument to measure ability is limited at $\theta \approx 0.5$, from a lack of sufficient items.

Relative Efficiency

The relative efficiency, which was introduced by Lord (1977), is the ratio of information functions from two tests (Lord, 1977). The relative efficiency can be a useful statistic for comparing the precision (i.e., reliability) of two tests regarding estimation of ability for specific levels of θ . This is especially important when selecting tests for evaluation of patients. For example, if a clinician wishes to assess function in athletes, they would desire an instrument that has increased precision when ability level θ is high. Relative efficiency can assist clinicians with determining which test is most appropriate for measuring function at the desired ability range.

Advantages of Item Response Theory over Classical Test Theory

Classical Test Theory (CTT) is a measurement model based upon tautology, that is, the mathematical model framework cannot be proven or falsified. Under CTT, the total score of an instrument indicates a patient's ability. Item parameters include item difficulty, which is the proportion of correct responses, and the item-total correlation, a measure of item discrimination. In general, the purpose of CTT is similar to IRT. Unlike IRT however, CTT is a sample-dependent framework (Hambleton, 1991; Lord, 1980). The difficulty of the items of an instrument are dependent upon the ability of those taking the test. For example, ability to climb stairs, a traditional item on instruments assessing activities of daily living, will have vastly different item difficulty estimates depending on whether the item is evaluated in elderly adults or athletes.

For knee-related function instruments, this is a major advantage of IRT. In IRT, item parameters are evaluated independently of patients, which provides invariance between multiple groups. This has important implications for measurement of function, where patients are expected to undergo large changes in function following injury and throughout treatment. After calibration procedures, items and patients are placed on the same scale. This will allow for the examination of the distribution of items across ability, the location of items, redundancy in items, and gaps in item-person distribution. This is vital for ensuring that the items of the instrument provide sufficient information about knee-related function across a wide range of ability. This information will help to identify areas in the ability distribution where additional items are needed. Figure 13 illustrates the advantage of placing persons and items on the same scale. It also depicts the spread of items and a missing gap between items where there is poor coverage (i.e., low discrimination).

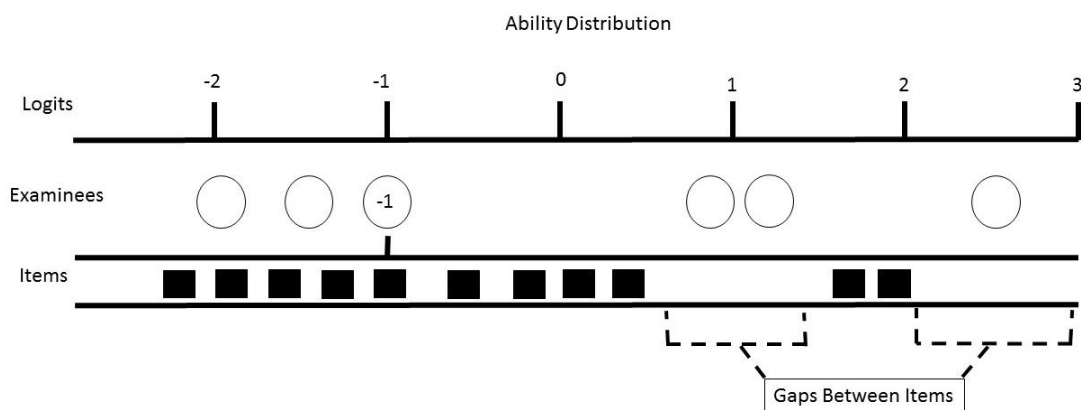


Figure 13. Illustration of Item-Parameters and Person-Ability Estimates Placed on the Same Scale. The ability measure for the person (white circle) to the far left was -1 and the item on the same measure directly under the person has a difficulty of -1. Distribution of items and gaps are also illustrated.

Historical Origins of Item Response Theory and Rasch Measurement Models

The preliminary foundations for IRT originated from Frederick Lord's investigation of raw-score distributions (Lord, 1953). Lord proposed that because two different tests of the same ability will generally not provide the same frequency distributions of raw scores; these raw score distributions are not reflective of the distribution of true ability (Lord, 1953). In addition, there were three technical reports presented to the United States Air Force by Birnbaum, which although not widely read helped to provide the foundations for IRT (A Birnbaum, 1957, 1958a, 1958b).

The information from Birnbaum's reports contributed to four chapters within Lord and Novick's book *Statistical Theories of Mental Test Scores*, which is often cited as one of the earliest references to IRT (Allan Birnbaum, 1968). This textbook was a pivotal piece of literature and well supported by many psychometricians. An ongoing educational testing services seminar, where Lord was a senior author, included many key measurement professionals and helped to popularize and establish the new latent trait models (Embretson & Reise, 2013).

Early work by Lord, Novick, and Birnbaum created interest in IRT and inspired Bock and several students to develop computer programs and estimation methods such as BILOG, TESTFACET, MULTILOG and PARSCALE. Bock along with Aitken were also responsible for developing the marginal maximum likelihood parameter estimation method, one of the more prominent estimation methods used today (Bock & Aitkin, 1981).

Although mathematically similar, there are key distinct philosophical differences between Rasch measurement theory and IRT. These two competing paradigms emerged around the same period. The Danish mathematician George Rasch developed the Rasch model to assist with test development for reading and the Danish military (Rasch, 1961).

Rasch later visited the United States where he shared his model and philosophies with Wright, a professor of education at the University of Chicago. Many of Wright's doctoral students (Andrich, Masters, Douglas and Wilson) (Andrich, 1978; G. N. Masters, 1982; Wright & Douglas, 1977; Wilson, 1992)

went on to provide significant contributions to Rasch measurement theory. Wright also conducted numerous lectures focused on objective measurement principles (Embretson & Reise, 2013).

Today both of these techniques are widely used and have many practical applications in a variety of fields. These models were introduced into Kinesiology in a series of papers published in *Research Quarterly for Exercise and Sport*; written by Safrit, Spray and Wood (Safrit, 1987; Spray, 1987; Wood, 1987). These papers introduced IRT using practical examples and encouraged researchers to begin adopting these modern testing models into their own research. More recent examples of the application of Rasch and IRT principles include the evaluation of the measurement properties of concussion tests (Ragan, Herrmann, Kang, & Mack, 2009), concussion symptom checklists (Chan, 2005), and development of item banks for PROMIS (Reeve et al., 2007).

Philosophical Differences between IRT and Rasch

Although the Rasch simple logistic model and the one-parameter IRT model are mathematically similar, there are distinct fundamental philosophical differences between these two paradigms (Andrich, 2004). These differences are important to discuss because of their implications in model selection and the interpretation of results. In both theories, the underlying principles and basic assumptions are relatively similar; the investigation of a patient's responses to items for determining the probability of correct responses.

When analyzing response patterns with IRT models, the goal is to identify the model that best fits the data. For example, a research may analyze response data using the one-, two-, and three-parameter models and evaluate which model provides the best fit. Measurement experts with IRT backgrounds believe that when response data does not fit a model well, the model is the problem and a new model is tested.

This philosophy is different from Rasch, where poor model-data fit is associated with bad items. Bad items are removed or revised and the data are re-assessed in a new sample to determine if the model has improved. For this reason, some consider Rasch to be a more appropriate model for test development, where the quality of the items are not yet established (Andrich, 2004).

Rasch Rating Scale Model

Analysis of ordered categorical data, such as those found on surveys and questionnaires, is not appropriate with the Rasch simple logistic model due to the sequential nature of the categories. Understanding which category a patient responded to has important clinical applications, particularly with monitoring changes in function over time. An alternative Rasch model, developed by David Andrich, known as the Rasch rating scale model is a psychometric model derived from Thurston's concept of thresholds used with ordered categorical data (Andrich, 1978).

Under the Rasch rating scale model when a patient responds to an item i , in one of m ordered categories ($j = 1, \dots, m_i$), the categories are separated by m

- 1 ordered thresholds on a continuum. A graph showing the relationship between the probability of responding in a given category as a function of person location is referred to as a Category Probability Curve (see figure 14). As person ability passes each threshold; the patient will be more likely to respond in successive categories. The Rasch rating scale model can be expressed mathematically with the following formula:

$$P_x(\theta) = \frac{e^{\sum_{j=0}^x [\theta - (\lambda_i + \delta_i)]}}{\sum_{x=0}^M e^{\sum_{j=0}^x [\theta - (\lambda_i + \delta_i)]}} \quad (12)$$

where (λ_i) is the scale location parameter, which reflects the relative difficulty of a particular item; and (δ_i) is the category intersection parameter, which *represents the number of* ordered thresholds, taken over all items in the instrument.

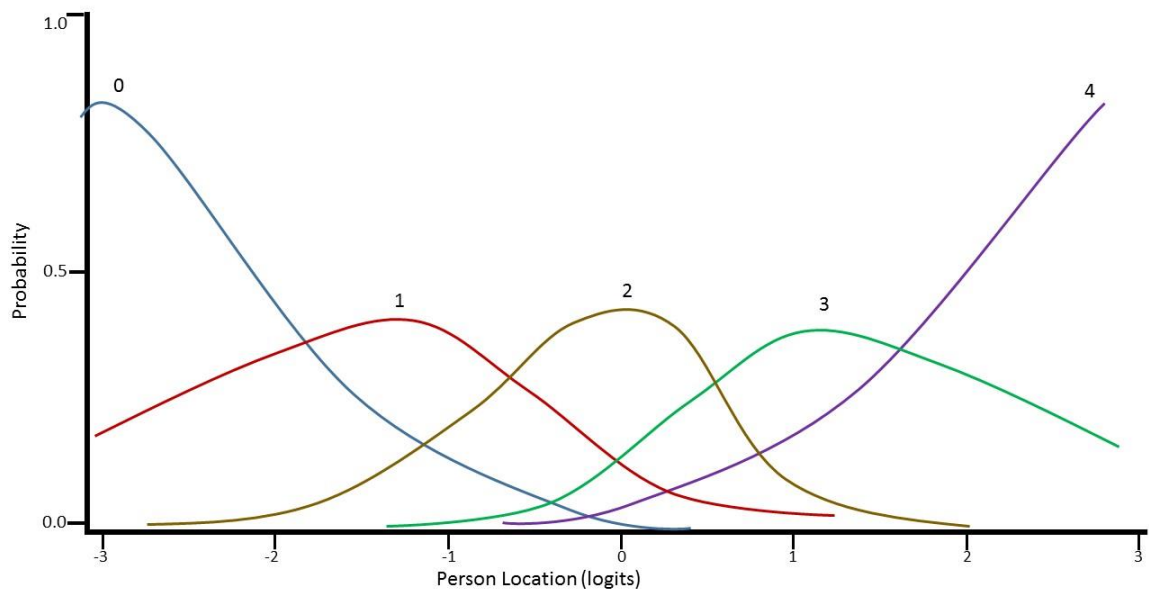


Figure 14. Rasch Rating Scale Model Category Probability Curve.

Rasch Partial Credit Model

Similar to the Rasch simple logistic model, the Rasch rating scale model assumes that all items have equal discrimination properties. The Rasch rating scale model, however, adds an additional assumption that all items have the same number of thresholds. In addition, the mean of the threshold locations is equivalent across all items. When the latter assumption is not met (i.e., the number of thresholds is not equivalent across all items), it is recommended that the Rasch partial credit model be used.

The Rasch partial credit model, developed by Greg Masters (G. N. Masters, 1982), is an extension of the Rasch rating scale model that allows for

differences in the number of thresholds across items. The Rasch partial credit model can be expressed mathematically with the following formula:

$$P_{ix}(\theta) = \frac{e^{\sum_{j=0}^x (\theta - \delta_{ij})}}{\sum_{x=0}^M e^{\sum_{j=0}^x (\theta - \delta_{ij})}} \quad (13)$$

where $\sum_{j=0}^x (\theta - \delta_{ij}) = 0$. The δ_{ij} term is sometimes called the item step difficulty associated with a category score of j ; the higher the value of a particular δ_{ij} , the more difficult a particular step is relative to other steps with an item. In other words, as the value of δ_{ij} increases it becomes less likely that a patient will respond to the higher categories. This model has important implications for use with functional measures. Currently many of the instruments use ordered-categorical-type questions where patients must report the degree/severity of the limitations association with knee-related injuries. In addition, the variety of questions included in these instruments necessitate a flexible model where each item can be evaluated independently.

Summary

There is no “gold standard” tool for evaluating knee-related function. Instruments should be developed following a strong theoretical framework, such as the WHO ICF. In addition, advanced measurement procedures such as Rasch measurement theory should be used to determine the quality of the items and their measurement properties. Many of the currently used instruments lack one, or both of these qualities. Furthermore, these instruments demonstrate significant

ceiling effects which may bias estimates of function. In athletic and high functioning individuals, these problems are likely to be more prevalent.

The application of Rasch measurement theory, particularly the Rasch partial credit model, has promise for improvement of functional measures. The distribution of items respective to ability level would provide evidence of the overall quality of these instruments. Because of the invariance property of Rasch, it would be possible to estimate how well these items perform at various ranges of function. It is expected, due to the ceiling effects that an insufficient number of items will be available to assess the higher levels of function.

The purpose of knee-related PRO is to measure function of patients throughout treatment. Because function is variable throughout treatment a variety of items are needed to measure function across the ability spectrum. Identification of gaps in item distribution would help to guide the development of new items. Moreover, item information functions will be computed for each item. In addition, the test information function will be used to assess the precision of knee-related PRO measures across the spectrum of ability. Therefore, the specific aims of this project are to:

Specific Aims

[1] To evaluate the psychometric properties of commonly used patient-reported knee-outcomes instruments in an athletic population in an effort to identify the gaps in item coverage of current instruments.

[2] To develop new items based upon content expert and patient review of psychometric data.

[3] To evaluate the new items in an athletic population to confirm improvement of existing instruments.

CHAPTER III

PYSCHOMETRIC EVALUATION OF PATIENT-REPORTED KNEE OUTCOMES FOR ASSESSMENT OF KNEE-RELATED FUNCTION

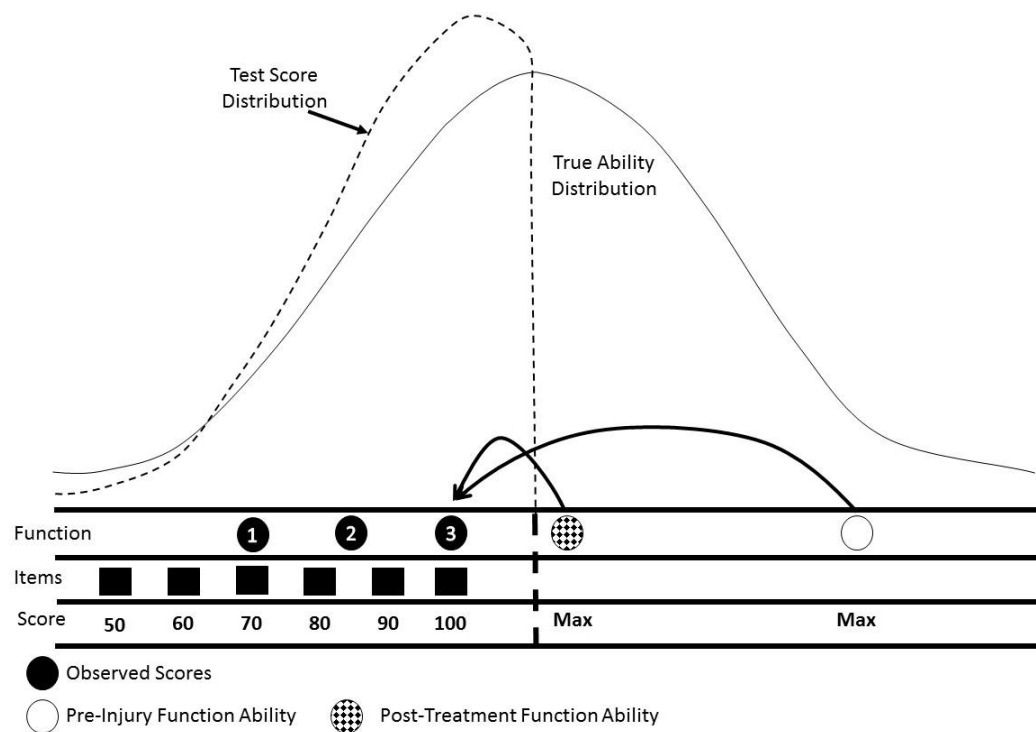
In 2015, the United States Physical Activity Council estimated that approximately 56% of the population (over the age of six) were involved in some form of sports-related activity (Physical Activity Council, 2016). Despite numerous health-related benefits, injuries are an unfortunate side effect of sports participation. Injuries in sports are problematic because they can negate the beneficial effects of sports participation, particularly when athletes are no longer able to participate due to residual effects of injuries (Engström et al., 1990; L. Lohmander, Östenberg, Englund, & Roos, 2004; Maffulli, Longo, Gougoulas, Loppini, & Denaro, 2010).

In intercollegiate sports, the knee joint is one of the most commonly injured regions of the body, accounting for approximately 15% of all injuries (Agel, Dick, et al., 2007; Agel, Dompier, et al., 2007; Agel, Evans, et al., 2007; Agel, Olson, et al., 2007; Agel, Palmieri-Smith, et al., 2007; Agel, Ransone, et al., 2007; Dick, Ferrara, et al., 2007; Dick, Hertel, et al., 2007; Dick, Lincoln, et al., 2007; Dick, Putukian, et al., 2007; Dick, Romani, et al., 2007; Dick, Sauers, et al., 2007; S. W. Marshall, Covassin, et al., 2007; S. W. Marshall, Hamstra-Wright, et al., 2007). In many cases these injuries resolve quickly with no long-term effects. When injuries are more severe, however, they require athletes to be removed from competition until they have sufficiently recovered from their injuries. The National Collegiate Athletic Association (NCAA) Injury Surveillance Program

(ISP) indicates that among all injuries resulting in extended lost time from sports competition (i.e., 10 or more days), up to 50% are related to the knee.

In order to ensure that athletes are receiving the most optimal care, appropriate documentation of the changes in health status, resulting from treatment, are required. A number of different patient-reported outcomes (PRO) measures have been developed to assess knee-related function following injury. The Knee Injury and Osteoarthritis Outcome Score (KOOS), International Knee Documentation Committee subjective knee form (IKDC-SKF), and the Marx Activity Rating Scale (MARS) are among the most popular and widely used knee-function instruments (Wang et al., 2010). These three instruments have been also recommended for use as PRO measures by the American Academy of Orthopaedic Surgeons (American Academy of Orthopaedic Surgeons, 2016).

Despite their use in clinical practice there are a variety of limitations associated with the use of these instruments. Specifically ceiling effects as high as high as 28% have been identified within some subscales of the KOOS. When ceiling effects are present in an instrument the ability of these instruments to accurately assess function is greatly diminished. An example of this phenomenon has been illustrated in figure 15.



*Figure 15. Theoretical Knee-Related Function Distribution. This figure illustrates a patient's knee-related function ability measured at three time points during recovery from an injury, which highlights a ceiling effect at the third measurement. Adapted from Ragan, BG (2004). *An empirical investigation of several critical psychometric issues in neuropsychological testing of mild traumatic brain injuries*. University of Illinois at Urbana-Champaign, Urbana, IL.*

In this example, a patient's knee-related function is evaluated multiple times following an injury to assess treatment efficiency and determine when it is safe to resume normal activity. As expected, the patient's knee-related function

score is low after injury (observed score 1) and improves over time as the injury heals (observed score 2 & observed score 3). After the third assessment; the instrument is no longer able to measure changes in function because there are no items capable of measuring the patient's function. This results in a lack of continuity in documentation of patient outcomes. Furthermore, this introduces the risk of returning patients to normal activity too soon, before their function has returned to pre-injury levels due to the inability of the instrument to appropriately document function. In athletic and high functioning individuals, these problems are likely to be more prevalent. Currently, however, there is limited data investigating the quality of these instruments in athletic populations where function is expected to be high.

The application of Rasch measurement theory, particularly the Rasch partial credit model, has promise for improvement of functional measures. Rasch modeling has the capability of assessing the quality of each item using advanced statistical measures as well as the performance of these items across a wide range of knee-related function. Identification of gaps in item distribution, through Rasch model analysis, could be used to help guide the development of new items. In addition, item and test information functions could be used to assess the precision of these instruments at various levels of knee function, which is necessary to determine the clinical usefulness of these instruments in athletic populations. Therefore, the purpose of this study is to investigate the psychometric properties of the KOOS, IKDC-SKF, and the MARS to better understand how these items perform in athletic populations.

Methods

Participants and Procedures

A convenience sample of 160 adults (mean age = 28.08 ± 10.95 ; male = 38.10%) were recruited at three separate universities and online using social media (Twitter, Facebook, etc.) for this cross-sectional study. Participant's activity levels ranged from highly active to inactive based upon the American College of Sports Medicine (ACSM) physical activity guidelines (Physical Activity Guidelines Advisory Committee, 2008). Because the emphasis of this study was to examine the performance of this instrument in highly functioning individuals, recruitment of athletes and active individuals was emphasized. Descriptive statistics for the participants of this study are provided in Table 5.

Table 5. Descriptive Statistics (n = 160).

Variable	$\bar{x}$	σ	n	%
Age	28.08	10.95		
Height (cm)	171.11	10.59		
Weight (kg)	78.60	19.77		
Sex				
Male			61	38.10
Female			99	61.90
Physical Activity Classification*				
Inactive			13	8.10
Lightly Active			47	29.40
Moderately Active			42	26.30
Highly Active			58	36.30
Sport Participation				
No			88	55.00
Yes			72	45.00
Current Knee Injury				
No			126	78.80
Yes			34	21.30

Abbreviations: $\bar{x}$ = mean; σ = standard deviation; n = number of participants; % = percentage of total sample.

*Activity Classifications: Inactive - May perform some light physical activity (such as walking, climbing stairs), but no moderate or vigorous physical activity); Lightly Active - Some moderate or vigorous intensity physical activity, but less than 150 minutes per week of moderate intensity physical activity, or 75 minutes of vigorous intensity physical activity; Moderately Active - 150 to 300 minutes per week of moderate intensity physical activity, 75 to 150 minutes of vigorous intensity physical activity, or a combination of the two; Highly Active - More than 300 minutes per week of moderate intensity physical activity, 150 minutes of vigorous intensity physical activity, or a combination of the two

Upon recruitment all participants were asked to complete a combined functional assessment questionnaire which contained all questions from the KOOS, the IKDC-SKF, the MARS and a brief injury history form. The questionnaire was hosted online using Google Forms and distributed through social media and email. Prior to beginning the survey participants were instructed to read and provide informed consent using the online consent document, which was approved by the Institutional Review Board from each institution where data was collected. To provide a more diverse range of ability estimates, the data collected was combined with existing pilot data from a previous study containing de-identified KOOS data from 79 adults with knee injuries resulting in a total sample size of 239 adults.

Instruments

Knee Injury and Osteoarthritis Outcome Score. The KOOS is a 42-item self-report PRO measure (Roos et al., 1998). The items within the KOOS include questions about pain (n = 9), symptoms (n = 7), activities of daily living (n = 17), function in sport and recreation (n = 5), and knee-related quality of life (n = 4). The score for each item ranges from 0 to 4 points. The total score for each domain is calculated by summing the score of each item within that domain. The summed score for each domain is then rescaled to a 100-point scale. For example, a patient with an item score 16 for the KOOS-pain subscale will have a total pain score of 44 ($16 * 100 / 36 = 44$). These scores represent a patient's level of knee function where a higher score indicates a better level of knee function.

International Knee Documentation Committee Subjective Knee Form. The IKDC-SKF is a 19-item self-report PRO measure (Irrgang et al., 2001). The items within the IKDC-SKF include questions about sports activities ($n = 10$), symptoms ($n = 7$), and function ($n = 2$). The score for the IKDC-SKF ranges from 0 to 100, with a higher score representing lower levels of symptoms and higher levels of function and sports activity.

Marx Activity Rating Scale. The MARS is a brief 4-item self-report PRO measure (Marx, Jones, et al., 2001). The 4-item scale evaluates how often an athlete performs running, cutting, deceleration, and pivoting activities. The score for each item ranges from 0 (less than one time in a month) to 4 (4 or more times in a week), where a higher score indicates a higher level of activity.

Data Analysis

Psychometric evaluation of the KOOS, IKDC-SKF, and the MARS occurred using the Rasch partial credit model. A Rasch model was estimated, which includes item difficulty (i.e., perceived difficulty of task, or severity of symptom), person ability parameters (i.e., the adult's level of knee-related function), and category function.

Item difficulty and person ability estimates are represented in logit values. Higher item difficulty logits represent a more challenging task, while higher person ability estimates represent an increased level of knee-related function. For polytomous items, such as those seen in Likert-type rating scales the likelihood of responding to a particular category for a given item is identified by

the boundary location (i.e., step difficulty). A separate item difficulty is reported for each boundary location; for example, a 5-category question will contain four item difficulties. The boundary locations for each item should be ordered.

The model-data fit for each item was evaluated using Infit and Outfit statistics. Infit and Outfit statistics are reported as mean-square residuals, which are chi-square statistics divided by their degrees of freedom, so that they have a ratio-scale form, with an expected value of one and range from zero to positive infinity. Mean-square residuals between 0.5 and 1.5 indicate that an item has acceptable fit (Wright, Linacre, Gustafson, & Martin-Lof, 1994). Values less than 0.5 indicate homogenization of scores, while values greater than 1.5 indicate large variability in scores (Wright et al., 1994).

Category function was evaluated for each item using the rating scale guidelines proposed by Linacre (Linacre, 2002). The item step difficulties for each item should increase for each category; each category should be endorsed by at least 10 participants; and step difficulties should advance by at least 1.4 logits, and less than 5.0 logits.

The test information function (TIF) and item information functions (IIF) were used to identify potential gaps in item distribution, across ability estimates. The TIF is a graphical representation of how much information the test is providing at each level of ability. To ensure reliable estimates of function, the TIF should contain high values across the entire range of ability for a given

population. The desired measurement precision for this instrument, as measured by the conditional standard error (inverse of the TIF) should be less than 0.3 at a given level of theta.

Results

Scoring summaries for each of the three PRO measures are listed in Table 6. Large ceiling effects were present in both the KOOS and the IKDC-SKF. On the IKDC-SKF and 3 subscales of the KOOS the maximum score was obtained at the 75th percentile. On the activities of daily living subscale of the KOOS, a score of 95 (out of 100 possible points) was associated with the 50th percentile.

Table 6. Patient Reported Outcomes – Knee Scoring Summary

Scale	n	25th Percentile	50th Percentile	75th Percentile
KOOS – Symptoms	239	57.14	78.57	92.86
KOOS – Pain	239	69.44	88.89	100.00
KOOS – ADLs	239	74.74	94.98	99.65
KOOS – Sport	235	50.00	80.00	100.00
KOOS – QOL	232	43.75	81.25	100.00
IKDC-SKF	160	79.31	94.25	100.00
MARS	160	1.00	8.00	12.00

Note. The maximum score for each subscale of the KOOS is 100 pts; the maximum score for the IKDC-SKF is 100 pts; the maximum score for the MARS is 16 pts. For each scale a higher score represents lower knee-related symptoms and higher knee function.

Abbreviations: KOOS = Knee Injury Osteoarthritis Outcomes Score; ADLs = Function, daily living; Sport = Function, sports and recreational activities; QOL = Quality of Life; IKDC-SKF = International Knee Documentation Committee Subjective Knee Form; MARS = Activity Rating Scale; n = sample size

Analysis of the three knee-related PRO measures with the Rasch partial-credit model indicated that many of the items had poor model-data fit. Of the 65 items evaluated in the original model, 35 had mean square residuals outside of the acceptable range ($0.5 \leq X \leq 1.5$) or at least one category with zero endorsement. To improve model-data fit the worst fitting item was removed and the model was re-analyzed with the remaining items. This process was repeated until all items had acceptable model-data fit. Test and item statistics for the final model have been provided in Table 7, which includes item step difficulties, person separation reliability, as well as the maximum item information function for each item and maximum test information function for each subscale. The item step difficulties for the remaining 30 items, ranged from -5.45 logits (least difficult) to 0.57 logits (most difficult). The person separation reliability is analogous to Cronbach's α . This is the degree to which the PRO measures differentiates persons in the test's outcome. Values range from 0 to 1.

Examination of category function suggests that the five-category response options used for the three knee-related PRO measures did not function well. The item step difficulties were disordered for 18 of the 30 items. Furthermore, 19 items had too few endorsements on one or more categories. Lastly, many of the items had minimal separation between categories.

Table 7. Rasch Partial-Credit Model Item Statistics for Final Model

Item	n	b ₀	b ₁	b ₂	b ₃	Infit	Outfit	Maximum IIF (θ)	Maximum TIF (θ)	Alpha
Full Test									-1.50	1.00
KOOS Symptom Scale									-0.75	0.87
KQ6	238	-4.58	-3.26	-0.90	-0.10	0.79	0.75	-0.75		
KQ7	236	-3.60	-0.97	-0.82	0.29	0.75	0.69	-0.75		
KOOS Pain Scale									-1.95	0.97
KQ8	232	-2.75	-1.80	-0.79	0.31	0.82	0.69	-1.35		
KQ9	232	-3.17	-3.87	-1.36	-3.73	0.82	0.64	-3.35		
KQ10	238	-2.66	-2.42	-1.25	-1.42	0.78	0.53	-1.85		
KQ11	238	-4.04	-3.95	-1.42	-1.23	1.04	0.84	-3.75		
KQ12	235	-2.89	-2.90	-1.60	-1.20	0.95	0.84	-2.30		
KQ13	235	-4.01	-2.84	-2.09	-0.32	1.03	0.85	-2.60		
KQ14	236	-4.01	-4.99	-2.06	-1.36	0.87	0.64	-4.00		
KQ15	237	-4.00	-5.01	-2.10	-1.14	0.81	0.64	-4.00		

Table 7. Rasch Partial-Credit Model Item Statistics for Final Model (cont.)

Item	n	b0	b1	b2	b3	Infit	Outfit	Maximum IIF (θ)	Maximum TIF (θ)	Alpha
KOOS Activities of Daily Living Scale									-2.80	0.98
KQ18	237	-3.88	-3.37	-1.66	-0.61	0.68	0.55	-3.15		
KQ19	236	-4.10	-3.60	-2.31	-0.65	0.77	0.58	-3.40		
KQ20	236	-4.21	-2.66	-1.85	-1.24	0.70	0.53	-2.05		
KQ21	236	-3.70	-3.53	-1.83	-1.02	0.81	0.65	-3.15		
KQ22	237	-4.56	-2.98	-1.72	-1.47	0.75	0.52	-2.00		
KQ30	237	-2.00	-5.45	-2.26	-1.46	0.79	0.62	-3.65		
KQ32	234	-2.80	-2.95	-1.39	-0.65	0.83	0.64	-2.45		
KOOS Sport/Recreation Scale									-1.55	1.00
KQ34	233	-1.95	-2.36	-0.74	0.14	0.80	0.67	-1.80		
KQ35	231	-1.88	-2.03	-0.98	-0.21	0.80	0.70	-1.55		
KQ36	226	-1.36	-2.19	-1.10	-0.31	0.72	0.56	-1.55		
KQ37	225	-1.89	-1.45	-1.75	-0.80	0.69	0.52	-1.55		

Table 7. Rasch Partial-Credit Model Item Statistics for Final Model (cont.)

Item	n	b ₀	b ₁	b ₂	b ₃	Infit	Outfit	Maximum IIF (θ)	Maximum TIF (θ)	Alpha
KQ38	223	-1.56	-1.99	-0.72	0.13	0.84	0.71	-1.45		
KOOS Quality of Life Scale									-1.30	0.96
KQ39	226	-1.72	-0.45	-0.52	0.57	0.71	0.62	-0.50		
KQ40	230	-1.78	-2.00	-0.87	-0.83	0.94	0.98	-1.40		
KQ41	229	-2.39	-1.69	-1.37	-0.86	0.93	0.64	-1.55		
IKDC-SKF									-0.95	0.89
IKDCQ1	160	-3.84	-0.91	-0.10	-0.94	0.93	0.66	-0.55		
IKDCQ5	160	-3.61	-1.27	-0.55	-1.20	0.98	0.70	-0.95		
IKDCQ7	160	-4.53	-0.48	-1.20	-1.39	0.91	0.69	-1.05		
IKDCQ9c	160	-1.97	-2.02	0.02	-0.41	0.91	0.71	-0.70		
IKDCQ9e*	160	-2.63	-1.54	-1.29	-	1.12	0.73	-1.70		

Note. b₀, b₁, b₂, & b₃ represent the item step difficulty between categories, where b₀ represents the boundary between the least extreme and second response category; while b₃ represents the boundary between the most extreme fourth response category.

*The two most extreme categories were collapsed due to zero endorsement of the most extreme category.

Abbreviations: n = sample size; IIF = item information function; TIF = test information function

The distributions of person-ability estimates (i.e., patient's knee function ability) and the TIF indicate that the distribution of items was inadequate for this sample of athletic adults (see figure 16). Desired precision for the instrument was set at a conditional standard error of less than 0.30 logits. The combined knee-related PRO measures had unacceptable precision as θ increased beyond 0.35 logits (see figure 17). The mean ability score of adults was 0.00 ± 1.58 ($M \pm SD$) logits.

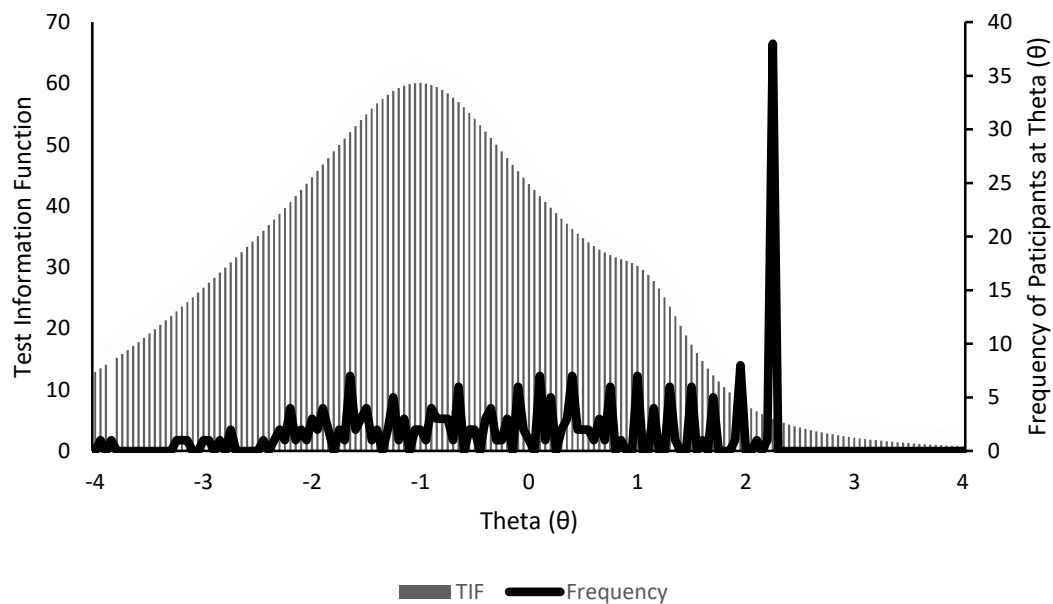


Figure 16. Test Information Function and Distribution of Person-Ability Estimates

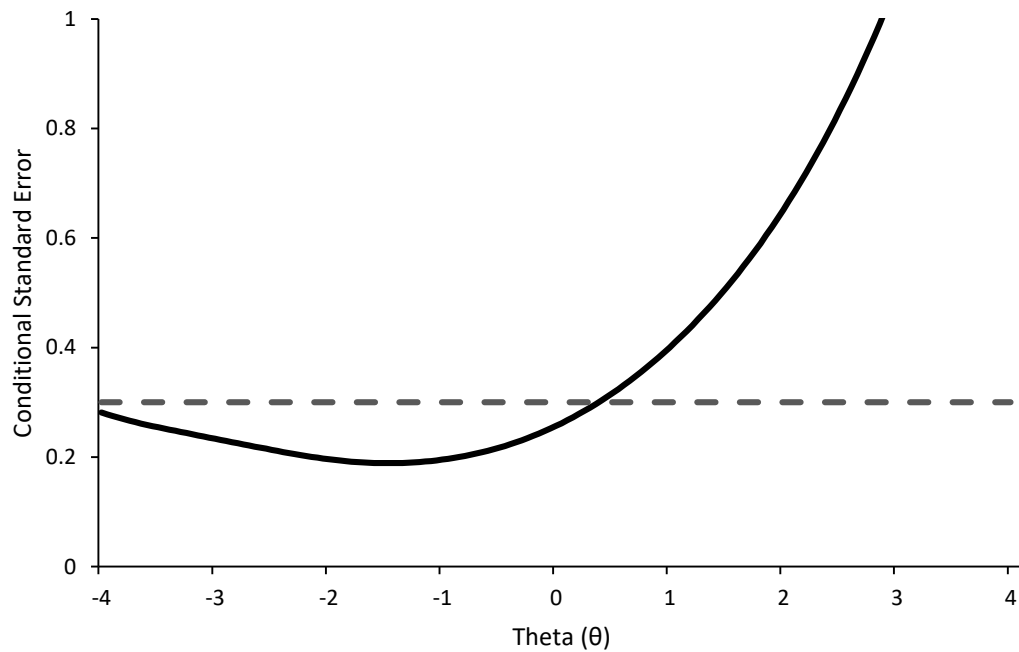


Figure 17. Conditional Standard Error of Combined Patient-Reported Knee Outcomes Measures. The dotted grey line represents the desired level of measurement precision (0.30 logits).

Discussion

The results from the Rasch calibration of the three knee-related PRO measures examined in this study suggests that although the precision of measurement may be acceptable for patients with low function ability, as function ability improves the measurement properties of these instruments are poor.

One of the major strengths of the Rasch analysis model is that the items and measurement precision of a test can be empirically examined and used to control and improve the quality of the test. Items that do not fit well can be identified and removed. In addition, because the items and participants are placed on a common metric, other items could be included on the same metric at a later time. From a test construction perspective, this provides test developers an opportunity to further shape and strengthen a measure.

Of the original 65 items contained in the combined functional instrument approximately 55% demonstrated poor model-data fit. Among the remaining items, 24 were from the KOOS, 5 from the IKDC-SKF, and 0 were retained from the MARS. The lack of items from the MARS is interesting considering that physical activity is a component of physical function (World Health Organization, 2001). Each of the MARS items assess general physical activity level of patients. While physical activity is a component of physical function, the lack of knee-specific information may be a limiting factor. Two of the items retained from the IKDC-SKF both assess activity level related to knee pain, and swelling. The inclusion of knee-specific information for the IKDC-SKF items may help explain the lack of model-data fit with the MARS items. Furthermore, studies have demonstrated that assessment of exercise capacity in adult patients suggests that physical function only accounts for approximately 26% of the variance in exercise METs (Coute, Ehrenfeld, Gupta, Terekhov, & Wanderer, 2017). These findings suggest that activity-related assessment may not be the best measure of physical function in this population.

Majority of the items with poor model-data fit were related to activities of daily living and symptoms. One possible explanation for this is the social stigma associated with pain and competition. Despite the negative consequences, many athletes often believe that “pain is a part of the game” (Deroche, Woodman, Stephan, Brewer, & Le Scanff, 2011). Particularly at the elite and college levels, many athletes will ignore pain and continue to participate at the same level of activity (Young, White, & McTeer, 1994). This may lead athletes to misreport their symptoms. Furthermore, studies have indicated that approximately half of patients achieve full knee-joint range of motion after only a month of rehabilitation post-surgery (van der List & DiFelice, 2017). Once a patient has achieved full range of motion, although there may still be some pain involved, they should be able to complete most basic activities of daily living. The low difficulty and challenge of these activities of daily living items contributes to the high ceiling effects of the instrument. More challenging items should be considered to help minimize and reduce this problem.

The primary goal of PRO measures should be to assess outcomes that are important to the individual. In most instances, the primary goal of an athlete following injury is to become healthy enough to return to play. In the activities of daily living subscale of the KOOS none of the items achieve this goal. This problem is seen with other PRO measures as well. For example, question 9 on the IKDC-SKF includes 9 tasks related to functional movement. Of these 9 tasks however, only four are actually sport related. The remaining 5 items (e.g., go up stairs, go down stairs, kneel on the front of your knee, sit with your knee bent,

and rise from a chair) all refer to activities of daily living. In this population these items may not be valuable to the patient, thus providing minimal useful information to the patient once they are able to complete these basic tasks.

The lack of challenging and meaningful items may also explain the low precision of these instruments at higher levels of function. Thirty-eight percent of participants demonstrated functional ability (person-ability range: -3.98 – 2.22 logits) that exceeded the item difficulty of the most challenging items. In addition, ceiling effects of at least 15% were identified in both the KOOS and IKDC-SKF. Because there are no items available at the higher level of function, estimation of function ability is biased. Low precision is a direct reflection of the reliability of an instrument. Minimal detectable change scores, which are determined from reliability, increase when accuracy is low. Minimal detectable change scores for KOOS subscales range from 14.3 – 19.6 for younger individuals and ≥ 20 points for older individuals (N. J. Collins et al., 2016). When change scores are high it is difficult to evaluate changes in patient function because such a large score is needed to determine true change. In order to estimate function more reliably, more information is needed about the patient.

Another major limitation of these measures is that many of the items had very few (or even zero) patients responding to the most extreme categories. For example, in the KOOS symptom subscale, the item, “How severe is your knee joint stiffness after first wakening in the morning?”, less than one percent of patients selected the response option “extreme” and only five percent of patients selected the option “severe”. Conversely, almost 50% of patients selected the

option “none”. Due to the lack of endorsement in the most extreme category it may be beneficial to reduce the number of category options for all items from 5 to 4, by combining the two most extreme categories.

Limitations

There were few limitations in this study. First, the data for this study relies on self-report information collected anonymously from participants online. Due to lack of physical contact with study participants there is no way to verify the validity of the participants’ injury claims. Furthermore, the analysis methods used in this study are complex and require large sample sizes. The de-identified injured dataset only contained responses for the KOOS instrument. As a result, the proportion of injured to healthy individuals is relatively small for the IKDC-SKF and MARS scales. The small sample size for these two scales may have influenced item-parameter estimates for these items; however, many of the items of the IKDC-SKF are duplicates or similar to the items of the KOOS.

Conclusion

Although the KOOS, IKDC-SKF and MARS are commonly used to monitor the recovery of knee-related function following injury in athletes. The data from this study suggests that as athletes recover from injury and their function levels improve the PRO measures evaluated in this study become less useful due to poor measurement properties. Future studies are needed to develop a knee-related functional scale that is suitable for measuring PRO in athletes throughout the rehabilitation process to ensure continuity of outcomes and increased precision at the higher ranges of function.

REFERENCES

- Agel, J., Dick, R., Nelson, B., Marshall, S. W., & Dompier, T. P. (2007).
Descriptive epidemiology of collegiate women's ice hockey injuries:
National Collegiate Athletic Association Injury Surveillance System, 2000-
2001 through 2003-2004. *Journal of Athletic Training, 42*(2), 249-254.
- Agel, J., Dompier, T. P., Dick, R., & Marshall, S. W. (2007). Descriptive
epidemiology of collegiate men's ice hockey injuries: National Collegiate
Athletic Association Injury Surveillance System, 1988–1989 through
2003–2004. *Journal of Athletic Training, 42*(2), 241-248.
- Agel, J., Evans, T. A., Dick, R., Putukian, M., & Marshall, S. W. (2007).
Descriptive epidemiology of collegiate men's soccer injuries: National
Collegiate Athletic Association Injury Surveillance System, 1988-1989
through 2002-2003. *Journal of Athletic Training, 42*(2), 270-277.
- Agel, J., Olson, D. E., Dick, R., Arendt, E. A., Marshall, S. W., & Sikka, R. S.
(2007). Descriptive epidemiology of collegiate women's basketball injuries:
National Collegiate Athletic Association Injury Surveillance System, 1988–
1989 through 2003–2004. *Journal of Athletic Training, 42*(2), 202-210.
- Agel, J., Palmieri-Smith, R. M., Dick, R., Wojtys, E. M., & Marshall, S. W. (2007).
Descriptive epidemiology of collegiate women's volleyball injuries: National
Collegiate Athletic Association Injury Surveillance System, 1988–1989
through 2003–2004. *Journal of Athletic Training, 42*(2), 295-302.
- Agel, J., Ransone, J., Dick, R., Oppliger, R., & Marshall, S. W. (2007).
Descriptive epidemiology of collegiate men's wrestling injuries: National

- Collegiate Athletic Association Injury Surveillance System, 1988-1989 through 2003-2004. *Journal of Athletic Training*, 42(2), 303-310.
- American Academy of Orthopaedic Surgeons. (2016). Patient Reported Outcome Measures. Retrieved from http://www.aaos.org/Quality/Performance_Measures/Patient_Reported_Outcome_Measures/
- Collins, N. J., Prinsen, C. A. C., Christensen, R., Bartels, E. M., Terwee, C. B., & Roos, E. M. (2016). Knee Injury and Osteoarthritis Outcome Score (KOOS): systematic review and meta-analysis of measurement properties. *Osteoarthritis and Cartilage*, 24(8), 1317-1329.
- Coute, R. A., Ehrenfeld, J. M., Gupta, D. K., Terekhov, M. A., & Wanderer, J. P. (2017). Electronically self-assessed functional capacity and exercise testing: A comparison of the Duke Activity Status Index and Patient-Reported Outcomes Measurement Information System tools. *American Heart Journal*, 188, 82-86.
- Deroche, T., Woodman, T., Stephan, Y., Brewer, B. W., & Le Scanff, C. (2011). Athletes' inclination to play through pain: a coping perspective. *Anxiety, Stress, & Coping*, 24(5), 579-587.
- Dick, R., Ferrara, M. S., Agel, J., Courson, R., Marshall, S. W., Hanley, M. J., & Reifsteck, F. (2007). Descriptive epidemiology of collegiate men's football injuries: National Collegiate Athletic Association Injury Surveillance System, 1988–1989 through 2003–2004. *Journal of Athletic Training*, 42(2), 221-233.

- Dick, R., Hertel, J., Agel, J., Grossman, J., & Marshall, S. W. (2007). Descriptive epidemiology of collegiate men's basketball injuries: National Collegiate Athletic Association Injury Surveillance System, 1988-1989 through 2003-2004. *Journal of Athletic Training, 42*(2), 194-201.
- Dick, R., Lincoln, A. E., Agel, J., Carter, E. A., Marshall, S. W., & Hinton, R. Y. (2007). Descriptive epidemiology of collegiate women's lacrosse injuries: National Collegiate Athletic Association Injury Surveillance System, 1988-1989 through 2003-2004. *Journal of Athletic Training, 42*(2), 262-269.
- Dick, R., Putukian, M., Agel, J., Evans, T. A., & Marshall, S. W. (2007). Descriptive epidemiology of collegiate women's soccer injuries: National Collegiate Athletic Association Injury Surveillance System, 1988-1989 through 2002-2003. *Journal of Athletic Training, 42*(2), 278-285.
- Dick, R., Romani, W. A., Agel, J., Case, J. G., & Marshall, S. W. (2007). Descriptive epidemiology of collegiate men's lacrosse injuries: National Collegiate Athletic Association Injury Surveillance System, 1988-1989 through 2003-2004. *Journal of Athletic Training, 42*(2), 255-261.
- Dick, R., Sauers, E. L., Agel, J., & Keuter, G. (2007). Descriptive epidemiology of collegiate men's baseball injuries: National Collegiate Athletic Association Injury Surveillance System, 1988-1989 through 2003-2004. *Journal of Athletic Training, 42*(2), 183-193.
- Engström, B., Forssblad, M., Johansson, C., & Tornkvist, H. (1990). Does a major knee injury definitely sideline an elite soccer player? *The American Journal of Sports Medicine, 18*(1), 101-105.

- Irrgang, J. J., Anderson, A. F., Boland, A. L., Harner, C. D., Kurosaka, M., Neyret, P., . . . Shelborne, K. D. (2001). Development and validation of the international knee documentation committee subjective knee form. *The American Journal of Sports Medicine*, 29(5), 600-613.
- Linacre, J. M. (2002). Optimizing rating scale category effectiveness. *Journal of Applied Measurement*, 3(1), 85-106.
- Lohmander, L., Östenberg, A., Englund, M., & Roos, H. (2004). High prevalence of knee osteoarthritis, pain, and functional limitations in female soccer players twelve years after anterior cruciate ligament injury. *Arthritis & Rheumatism*, 50(10), 3145-3152.
- Maffulli, N., Longo, U. G., Gougoulas, N., Loppini, M., & Denaro, V. (2010). Long-term health outcomes of youth sports injuries. *British Journal of Sports Medicine*, 44(1), 21-25.
- Marshall, S. W., Covassin, T., Dick, R., Nassar, L. G., & Agel, J. (2007). Descriptive epidemiology of collegiate women's gymnastics injuries: National Collegiate Athletic Association Injury Surveillance System, 1988-1989 through 2003-2004. *Journal of Athletic Training*, 42(2), 234-240.
- Marshall, S. W., Hamstra-Wright, K. L., Dick, R., Grove, K. A., & Agel, J. (2007). Descriptive epidemiology of collegiate women's softball injuries: National Collegiate Athletic Association Injury Surveillance System, 1988–1989 through 2003–2004. *Journal of Athletic Training*, 42(2), 286-294.
- Marx, R. G., Jones, E. C., Allen, A. A., Altchek, D. W., O'Brien, S. J., Rodeo, S. A., . . . Wickiewicz, T. L. (2001). Reliability, validity, and responsiveness of

- four knee outcome scales for athletic patients. *Journal of Bone and Joint Surgery American Volume*, 83(10), 1459-1469.
- Physical Activity Council. (2016). *The Physical Activity Council's annual study tracking sports, fitness, and recreation participation in the US*.
- Physical Activity Guidelines Advisory Committee. (2008). Physical activity guidelines advisory committee report, 2008. *Washington, DC: US Department of Health and Human Services, 2008, A1-H14*.
- Roos, E. M., Roos, H. P., Lohmander, L. S., Ekdahl, C., & Beynnon, B. D. (1998). Knee Injury and Osteoarthritis Outcome Score (KOOS)—development of a self-administered outcome measure. *Journal of Orthopaedic & Sports Physical Therapy*, 28(2), 88-96.
- van der List, J. P., & DiFelice, G. S. (2017). Range of motion and complications following primary repair versus reconstruction of the anterior cruciate ligament. *Knee*, 24(4), 798-807.
- Wang, D., Jones, M. H., Khair, M. M., & Miniaci, A. (2010). Patient-reported outcome measures for the knee. *The Journal of Knee Surgery*, 23(3), 137-152.
- World Health Organization. (2001). *International Classification of Functioning, Disability and Health*: Geneva: World Health Organization.
- Wright, B. D., Linacre, J. M., Gustafson, J., & Martin-Lof, P. (1994). Reasonable mean-square fit values. *Rasch Measurement Transactions*, 8(3), 370.
- Young, K., White, P., & McTeer, W. (1994). Body talk: male athletes reflect on sport, injury, and pain. *Sociology of Sport Journal*, 11(2), 175-194.

CHAPTER IV

DEVELOPMENT AND VALIDATION OF THE PATIENT REPORTED KNEE OUTCOMES ASSESSMENT TOOL (PRKOAT)

Injuries are an unfortunate side-effect of participation in sport with approximately 3.7 million individuals reporting to emergency rooms for sport-related injuries each year (Burt & Overpeck, 2001; Gage et al., 2012). Among these injuries, knee injuries are the second most common with reported incidence rates of approximately 600,000 injuries per year (Gage et al., 2012). Proper treatment and evaluation of knee injuries is important because of the risk of lasting long-term consequences. Studies have shown that knee injuries are associated with impairments in muscle function, knee pain, decreased knee function, and diminished quality of life (Flosadottir, Frobell, Roos, & Ageberg, 2017). In addition, many patients struggle to achieve pre-injury levels of function with knee injuries requiring surgical intervention, despite numerous improvements in surgical techniques and medical advancements (Aglietti, Giron, Buzzi, Biddau, & Sasso, 2004; Gobbi & Francisco, 2006; Kvist, Ek, Sporrstedt, & Good, 2005; Lee, Karim, & Chang, 2008; Wiger, Brandsson, Kartus, Eriksson, & Karlsson, 1999). In order to continue to improve treatment outcomes for patients as well as ensuring that patients are receiving the most optimal care, appropriate documentation of patient-reported outcomes (PRO) is critically important.

In the clinical setting, PRO measures provide numerous advantages for both clinicians and patients. More specifically, PRO measures enhance treatment adherence and improve communication between the patient and their clinician

through the measurement of important subjective outcomes such as quality of life and other psychological, sociological, and physiological factors that are specific in the individual (Institute of Medicine [US] Committee on Quality of Health Care in America, 2001). Examples of commonly used PRO measures include the Knee Injury and Osteoarthritis Outcomes Score (KOOS) and the International Knee Documentation Committee's Subjective Knee Form (IKDC-SKF).

In the general population the KOOS has been reported to yield adequate levels of content validity, internal consistency, test-retest reliability, construct validity, and responsiveness when used to assess knee-related function (N. J. Collins et al., 2016). Because of the intense demands of sports participation, however, athletes are a particularly unique population. When compared with non-athletes, athletes routinely score higher on PRO measures (Snyder et al., 2010). In a sample of 41 healthy-college-aged students with previous history of knee injury 100% of participants obtained the maximum score on the IKDC-SKF. When the maximum score is obtained on PRO measures these instruments are no longer able to monitor changes in patient function. This phenomenon creates a discontinuity in documentation of patient outcomes as clinicians are forced to seek alternative measurement options. Despite these limitations these instruments are commonly used in clinical practice. Unfortunately, few PRO measures used in sports are actually developed or validated using athletes (Irrgang et al., 2001; Johanson, Liang, Daltroy, Rudicel, & Richmond, 2004; Roos et al., 1998).

Rasch partial-credit model calibration of the KOOS, IKDC-SKF, and Marx Activity Rating Scale indicated that model-data fit was poor in all but 30 of the 65 items (Farnsworth II, Evans, Binkley, & Kang, n.d.). In addition, measurement precision was relatively low at higher levels of function. Furthermore, many of the items also exhibited poor category function. To improve functional assessment of athletes a new PRO measure, the Patient Reported Knee Outcomes Assessment Tool (PRKOAT), was developed.

The PRKOAT included 18 items from the KOOS, 5 items from the IKDC-SKF, 7 revised items, 9 new items, and 3 open-ended items for a total of 42 unique items. Although the previous study identified 30 acceptable items, some of these items were too similar or duplicate items contained in multiple instruments. Thus, 7 of the original items were modified to improve clarity, combine duplicate and similar items, make the items more personable, or to split items into multiple items. For example, a previous item “Difficulty with running” was separated into two unique items, “Difficulty with jogging (50% intensity)” and “Difficulty with running at maximum speed”. Because running is more challenging than jogging this provides an opportunity to validate the item hierarchy as well as providing more challenging and difficult items for student-athletes.

The 9 new items were developed to increase the difficulty of the test and minimize potential ceiling effects associated with low item difficulty. An example of one of the new items includes: “Difficulty with performing a single leg squat on your injured knee”. This particular item was included because it was believed to be a more challenging item than current functional items. Furthermore, many

ACL rehabilitation programs often include single leg strengthening and balance-related exercises to minimize limb asymmetry (Paterno, Ford, Myer, Heyl, & Hewett, 2007). Studies indicate that many ACL-deficient patients cannot perform stable single leg squatting, with evidence suggesting biomechanical differences between athletes before and after their injury (Yamazaki, Muneta, Ju, & Sekiya, 2009). Thus, this item should have a higher item difficulty and provide a better range of assessment in higher functioning individuals.

This new instrument has the potential to improve assessment of function in athletes. Rasch calibration of this new instrument is required, however, to determine the measurement properties of this instrument. Therefore, the purpose of this study is to examine the measurement properties of the newly developed PRKOAT in a sample of student-athletes.

Methods

Evaluation of the PRKOAT occurred in two distinct phases. The first phase of the study was used to pilot test the newly developed PRKOAT. Following completion of the pilot test, items were revised and re-assessed during the second phase of the study. Data collection for both phases of this study were approved by the Institutional Review Board at the university of the primary investigator.

Phase I: Pilot Test

Participants. A convenience sample of 32 Division III student-athletes (mean age = 20.78 ± 1.01 ; males = 56.30%) were recruited using flyers and word of mouth. Interested participants were instructed to email the primary investigator

for more information. Upon recruitment, participants were directed to the online survey, which was hosted using Google Forms. Upon accessing the survey participants read a brief explanation of the study and provided consent for completing the study by clicking “next” and beginning the survey.

Data Analysis. Psychometric evaluation of the PRKOAT occurred using the Rasch partial-credit model; which includes item difficulty (i.e., perceived difficulty of task, or severity of symptom), person ability parameters (i.e., the student-athlete’s level of knee-related function), and category function.

Item difficulty and person ability estimates are represented in logit values. Higher item difficulty logits represent a more challenging task, while higher person ability estimates represent an increased level of knee-related function. For polytomous items, such as those seen in Likert-type rating scales the likelihood of responding to a particular category for a given item is identified by the boundary location (i.e., step difficulty). A separate item difficulty is reported for each boundary location; for example, a 5-category question will contain four item difficulties. The boundary locations for each item should be ordered.

The model-data fit for each item was evaluated using Infit and Outfit statistics. Infit and Outfit statistics are reported as mean-square residuals, which are chi-square statistics divided by their degrees of freedom, so that they have a ratio-scale form, with an expected value of one and range from zero to positive infinity. Mean-square residuals between 0.5 and 1.5 indicate that an item has acceptable fit (Wright et al., 1994). Values less than 0.5 indicate homogenization

of scores, while values greater than 1.5 indicate large variability in scores (Wright et al., 1994).

Item Evaluation & Instrument Modification. Rasch partial-credit model analysis of the 39 pilot-test items (the 3 open-ended items were not included in the Rasch model analysis) indicated that 25 items had acceptable model-data fit. Item statistics for the Rasch model analysis have been provided in table 8. Twenty-four of the original items were either revised or deleted from the pilot test based upon results of the Rasch model analysis and participant feedback. In addition, two new items were developed. A summary of the item modifications and rationale has been provided in table 9.

Table 8. Item Statistics – Pilot Test (n = 32)

Item	Infit	Outfit	b ₀	b ₁	b ₂
1. How severe is your knee joint stiffness?	0.955	0.684	-1.162	-3.627	-
2. How severe is your knee joint stiffness after sitting, lying, or resting later in the day?	0.726	0.642	0.011	-	-
3. How often do you experience knee pain?	0.594	0.600	1.178	-0.543	-2.975
4. Pain with twisting/pivoting on your knee?	1.040	0.901	-1.090	-	-
5. Pain with straightening knee fully?	1.047	1.445	-1.506	-	-
6. Pain with bending knee fully?	0.986	0.572	-2.253	-	-
7. Pain with walking on a flat surface?	0.962	0.418	-3.434	-	-
8. Pain with going up or down stairs?	1.096	1.034	-1.854	-3.323	-
9. Pain at night while in bed?	0.912	0.432	-2.942	-	-
10. Pain with sitting or lying?	1.021	0.782	-2.253	-	-
11. Difficulty with ascending/descending stairs?	0.822	0.671	-3.312	-2.362	-
12. Difficulty with rising from sitting?	0.843	0.425	-2.819	-2.733	-
13. Difficulty with getting in/out of your vehicle?	0.787	0.252	-3.434	-	-
14. Difficulty with standing?	0.963	0.416	-2.819	-2.733	-
15. Difficulty with bending to the floor?	0.701	0.335	-3.175	-2.128	-

Table 8. Item Statistics – Pilot Test (n = 32) (cont.)

Item	Infit	Outfit	b ₀	b ₁	b ₂
16. Difficulty with picking up a small object (e.g., penny) off the floor?	1.092	0.665	-4.213	-	-
17. Difficulty with lifting a large object off the floor?	0.981	0.514	-2.903	-	-
18. Difficulty with walking on a flat surface?	0.971	0.438	-3.434	-	-
19. Difficulty with sitting with your knee bent?	0.895	0.558	-1.606	-3.443	-
20. Difficulty with performing double leg squats (body weight only?)	1.066	0.758	-2.566	-	-
21. Difficulty with performing single leg squats on your injured knee (body weight only)?	0.543	0.400	-1.452	-1.950	-2.921
22. Difficulty with lower body resistance training (e.g., weight lunges, weighted squats)?	0.880	0.626	-1.246	-2.770	-
23. Difficulty with jogging (approximately 50% intensity)	1.206	0.711	-2.253	-	-
24. Difficulty with running at maximum speed?	1.286	0.862	-2.002	-2.402	-
25. Difficulty with quickly changing direction while running (e.g., agility drills)?	0.666	0.509	-0.823	-2.926	-
26. Difficulty with jumping forward and landing on your injured knee?	1.020	0.617	-1.745	-1.549	-
27. Difficulty with jumping to the side and landing on your injured knee?	0.732	0.522	-1.293	-	-
28. Difficulty with hopping repeated on your injured knee?	0.729	0.370	-1.854	-3.323	-
29. Difficulty with twisting/pivoting on your injured knee?	0.589	0.436	-1.030	-2.854	-

Table 8. Item Statistics – Pilot Test (n = 32) (cont.)

Item	Infit	Outfit	b ₀	b ₁	b ₂
30. Difficulty with kneeling on the front of your injured knee?	1.389	1.327	-2.063	-1.361	-
31. How often are you aware of your knee problems?	1.020	0.893	-0.075	-1.021	-1.363
32. Have you modified your lifestyle to avoid potentially damaging activities?	1.602	2.056	-0.890	-1.061	-
33. How much are you troubled with lack of confidence in your knee?	0.876	0.453	-1.874	-1.917	-
34. What is the highest level of activity that you can perform without significant knee pain?	0.869	0.404	-1.960	-2.695	-
35. What is the highest level of activity that you can perform without significant swelling in your knee?	1.080	0.943	-2.253	-	-
36. What is the highest level of activity that you can perform without significant giving way in your knee?	0.907	0.325	-3.434	-	-
37. Compared to your non-injured peers (e.g., teammates or friends of similar athletic ability) how would you rate your overall knee function?	0.861	0.838	0.751	-1.745	n/a
38. Compared to your healthy knee, how would you rate your injured knee's overall level of function?	1.022	1.046	1.646	-2.147	n/a
39. Compared to your non-injured peers (e.g., teammates or friends of similar athletic ability) how would you rate your usual level of physical activity?	1.153	1.211	0.906	-	n/a

Note. Category options with zero responses were merged using XCalibre v.4.0. Shaded rows represent misfitting items or items with category dysfunction.

Table 9. PRKOAT Pilot Test Modifications

Item	Modification	Rational
7. Pain with walking on a flat surface?	Item Deleted	Item was too easy (30/32 in least severe category) and poor model data fit. Although the item is related to ADL the item is not sport-related. As patient progresses through rehab this item will likely become irrelevant early in the rehabilitation process.
9. Pain at night while in bed?	Item Deleted	Item was too easy (29/32 in least severe category) and poor model data fit. Although the item is related to ADL the item is not sport-related. As patient progresses through rehab this item will likely become irrelevant early in the rehabilitation process.
11. Difficulty with ascending/descending stairs?	Item Deleted	Item is almost identical to item #8 (pain with going up or down stairs) and the language between the two was inconsistent. Participant feedback suggested that pain-related items would be more relevant to athletes because many athletes often perform activities despite pain due to high mental toughness.
12. Difficulty with rising from sitting?	Item Revised → <i>Pain while rising from an armless chair?</i>	Participants reported that the lack of clarity related to the type of chair could lead to some confusion. The degree of difficulty with rising from an armless chair is vastly different than one with arm rest, which can be used to assist with standing. In the original item no participants responded to the most extreme category. Item changed from difficulty to pain based upon participant feedback.

Table 9. PRKOAT Pilot Test Modifications (cont.)

Item	Modification	Rational
13. Difficulty getting in/out of your vehicle?	Item Deleted	Item is too easy (30/32 in least severe category) and poor model data fit. Although the item is related to ADL the item is not sport-related. As patient progresses through rehab this item will likely become irrelevant early in the rehabilitation process.
14. Difficulty with standing?	Item Revised → <i>Pain while standing for an extended period of time (minimum 1 hour)?</i>	Participants reported that the lack of time frame associated with item could lead to ambiguity. Item revised to include a time frame. Time frame of at least one hour was selected in an effort to increase difficulty of item. Item changed from difficult to pain based upon participant feedback.
15. Difficulty with bending to the floor?	Item Deleted	Item had poor model-data fit. Additionally, it can be argued that bending to the floor is not directly related to the knee; bending is a hip-related movement which may or may not require the knee.
16. Difficulty with picking up a small object (e.g., penny) off the floor?	Item Revised → <i>Pain while picking up a small object (e.g., penny) off the floor?</i>	Item changed from difficulty to pain based upon participant feedback.
17. Difficulty with lifting a large object off the floor?	Item Revised → <i>Pain while lifting a large object off the floor?</i>	Item changed from difficulty to pain based upon participant feedback.

Table 9. PRKOAT Pilot Test Modifications (cont.)

Item	Modification	Rational
18. Difficulty with walking on a flat surface?	Item Deleted	Item was too easy (31/32 in least severe category) and poor model data fit. Although the item is related to ADL the item is not sport-related. As patient progresses through rehab this item will likely become irrelevant early in the rehabilitation process.
19. Difficult with sitting with your knee bent.	Item Revised → <i>Pain while sitting with your knee bent (at any degree) for an extended period of time (minimum 1 hour)?</i>	Participants reported that the lack of time frame associated with item could lead to ambiguity. Item revised to include a time frame. Time frame of at least one hour was selected because of the length of a typical college class. Item changed from difficulty to pain based upon participant feedback.
20. Difficulty with performing double-leg squats (body weight only)?	Item Revised → <i>Pain while performing standard squats (body weight only)?</i>	Participants felt that the use of the term “standard” was more appropriate “double leg” squats. Item changed from difficulty to pain based upon participant feedback.
21. Difficulty with performing single-leg squats (body weight only) on your injured knee?	Item Revised → <i>Pain while performing single-leg squats (body weight only) on your injured leg?</i>	Item changed from difficulty to pain based upon participant feedback. Changed the word “knee” to “leg” to maintain consistency with other items.

Table 9. PRKOAT Pilot Test Modifications (cont.)

Item	Modification	Rational
22. Difficulty with lower body resistance training (e.g., weight lunges, weighted squats)?	Item Revised → <i>Pain while performing lower body resistance training (e.g., lunges, step-ups, or squats)?</i>	Item changed from difficulty to pain based upon participant feedback. Removed the terms “weight” and “weighted” from examples because of redundancy and to shorten item.
23. Difficulty with jogging (approximately 50% intensity)	Item Revised → <i>Pain while jogging (approximately 50% intensity)</i>	Item changed from difficulty to pain based upon participant feedback.
24. Difficulty with running at maximum speed?	Item Revised → <i>Pain while running at maximum speed?</i>	Item changed from difficulty to pain based upon participant feedback.
Developed New Item	<i>Pain while back pedaling (i.e., running backwards)</i>	New item representing commonly used sport specific skill. A component of many agility tests.
25. Difficulty with quickly changing direction while running (e.g., agility drills)?	Item Revised → <i>Pain associated with changing direction while running (e.g., agility drills)?</i>	Item changed from difficulty to pain based upon participant feedback.
26. Difficulty with jumping forward and landing on your injured knee?	Item Revised → <i>Pain with jumping forward and landing on your injured knee?</i>	Item changed from difficulty to pain based upon participant feedback.

Table 9. PRKOAT Pilot Test Modifications (cont.)

Item	Modification	Rational
27. Difficult with jumping to the side and landing on your injured knee?	Item Revised → <i>Pain with jumping to the side and landing on your injured knee?</i>	Item changed from difficulty to pain based upon participant feedback.
28. Difficulty with hopping repeatedly on your injured knee?	Item Revised → <i>Pain with repeated high-impact jumps in place (such as jump knee tucks)?</i>	Item had poor model-data fit. Increased item difficulty by changing intensity of jump in an effort to improve model-data fit and increase test difficulty.
29. Difficulty with twisting/pivoting on your injured knee?	Item Deleted	Item is almost identical to item #4 (pain with twisting/pivoting on your injured knee). Participant feedback suggested that pain-related items would be more relevant to athletes because many athletes often perform activities despite pain due to high mental toughness.
30. Difficulty with kneeling on your injured knee?	Item Revised → <i>Pain with kneeling on your injured knee?</i>	Item changed from difficulty to pain based upon participant feedback.
32. Have you modified your lifestyle to avoid potentially damaging activities?	Item Revised → Have you modified your daily exercise or practice routines to avoid painful or potentially damaging activities?	Item had poor model data fit (too much variability). Revised item to be more specific to athletic population.

Table 9. PRKOAT Pilot Test Modifications (cont.)

Item	Modification	Rational
33. How much are you troubled with lack of confidence in your knee?	Item Revised → <i>To what extent are you troubled with lack of confidence in your knee when engaging in intense physical activity (such as during practice or games)?</i>	Item had poor model data fit. Revised item to be more specific to athletic population.
New Item	<i>To what extent do you feel anxious about performing certain activities because of your injured knee?</i>	Included question related to sport psychology based upon participant feedback.
34. What is the highest level of activity that you can perform without significant knee pain?	Item Deleted	Item had poor model-data fit.
36. What is the highest level of activity that you can perform without significant giving way in your knee?	Item Deleted	Item had poor model-data fit.

Many of the items in the pilot test had few or zero responses in the most extreme category; most commonly occurring with items assessing difficulty associated with performing basic tasks of daily living. In addition, the step difficulty (item category threshold between adjacent categories) was low between the middle categories for some items. In an effort to improve category function all of the items with five-category response options were revised to four-category response options. By combining the middle categories, the thresholds between categories should improve allowing for better distinction between individuals responding in each category.

An interesting observation was that items related to pain often performed superior to those related to difficulty. Sport psychology and sociology research reports that many athletes see “pain” as just a natural part of sports (Young et al., 1994). This suggests that many athletes may perform a challenging task and report “no difficulty” in spite of pain. Thus, for athletes, being able to complete a task may be less relevant than the pain associated with completing a task. Therefore, in this study items associated with “difficulty” were modified to “pain”.

Following all revisions, the revised PRKOAT contained 32 items (see Appendix I). Flesh-Kincaid reading analysis suggests the revised instrument requires a 10th grade reading level, which is appropriate for college-aged athletes (Kincaid, Fishburne Jr, Rogers, & Chissom, 1975). Scoring for items with a 4-point Likert scale ranged from 0 to 3; while scoring for 3-point Likert scale items ranged from 0 to 2. The maximum score (extreme functional limitations) for the PRKOAT will be a 78. A score of 0 will indicate no functional limitations.

Phase 2: Rasch Calibration

Participants. A convenience sample including 203 student-athletes (mean age = 21.46 ± 4.64 ; males = 54.70%) were recruited for calibration of the PRKOAT from three separate institutions, and online using social media (e.g., Twitter, Facebook). Majority of participants (i.e., 77.3%) self-reported as college-level division III athletes. Demographic information for all participants has been provided in Table 10.

The data collection methods for the study were identical to those used in the pilot study. Prior to accessing the online study, participants, were asked to read an informed consent, explaining the procedures and purpose of the study. Consent was obtained once participants clicked “next” after reviewing the consent information on the main survey page. In an effort to increase recruitment and compliance, participants in the study were offered the opportunity to win one of four \$25 Visa gift cards to be dispersed at the conclusion of the study. Participants who provided answers to all appropriate questions were entered into a drawing and randomly selected.

Table 10. Participant Demographic Information – Phase 2 (n = 203).

Variable	$\bar{x}$	σ	n	%
Age	21.46	4.64		
Height (cm)	175.76	11.06		
Weight (kg)	79.85	18.99		
Sex				
Male			111	54.68
Female			90	44.33
Prefer not to say			2	0.99
Sport Competition Level				
College – Division I			18	8.90
College – Division II			1	0.50
College – Division III			157	77.30
College – Club Sports			2	1.00
High School			1	0.50
Professional			2	1.00
Did not report			22	10.80
Sport				
Baseball			27	13.30
Basketball			11	5.42
Track & Field			26	12.81
Football			32	15.76
Wrestling			9	4.43
Rugby			1	0.49
Golf			12	5.91
Tennis			10	4.93
Soccer			23	11.33
Softball			13	6.40
Volleyball			6	2.96
Did not report			33	16.26
Current Injury				
Yes			36	17.70
No			167	82.30

Data Analysis. Psychometric evaluation of the PRKOAT followed a similar procedure as the pilot test analysis using the Rasch partial-credit model. Evidence of known-group difference validity was established through comparison of raw scores between injured and non-injured athletes. Significance was set at alpha 0.05. In addition, test information function (TIF) and item information functions (IIF) were used to identify potential gaps in item distribution, across ability estimates. The TIF is a graphical representation of how much information the test is providing at each level of ability. To ensure reliable estimates of function, the TIF should contain high values across the entire range of ability for a given population.

Category function was evaluated for each item using the rating scale guidelines proposed by Linacre (Linacre, 2002). The item step difficulties for each item should increase for each category; each category should be endorsed by at least 10 participants; and step difficulties should advance by at least 1.4 logits, and less than 5.0 logits.

Results

Overall the data fit the model well. Of the original 32 items examined on the PRKOAT, 27 had acceptable infit and outfit statistics. The 5 misfitting items were each removed one at a time until all items had acceptable model-data fit. The eliminated items have been listed in table 11.

Table 11. Misfitting Items

Item
Pain while rising from an armless chair?
Pain associated with changing direction while running (e.g., agility drills)?
Have you modified your daily exercise or practice routines to avoid painful or potentially damaging activities?
Pain with repeated high-impact jumps in place (such as jump knee tucks)?
Pain with jumping to the side and landing on your injured knee?

The item parameters of the PRKOAT and the ability parameters of examinees' knee function score were calibrated and placed on a common logit metric, which allowed the examination of the relative positions of items and examinees. A lower logit value indicates lower levels of knee function. Item difficulties (i.e., knee function) ranged from -4.74 to 1.89 logits. Person ability estimates ranged from -3.24 to 2.29 (mean 0.00 ± 1.35). Cronbach's α was 0.977, which represents the reproducibility of the raw scores identified in this study. Values range from 0 to 1.

The five least difficult items were [1] pain with straightening of knee fully, [2] knee joint stiffness severity, [3] pain while lifting a large object off the floor, [4] pain while jogging at 50% intensity, and [5] pain with bending knee fully. The five most difficult items were [1] how often are you aware of your knee problems, [2]

pain with kneeling on your injured knee, [3] compared to your non-injured peers
 how would you rate your overall knee function, [4] compared to your healthy knee
 how would you rate your knee's overall level of function, and [5] pain while
 performing single leg squats on your injured leg (body weight only). Detailed item
 statistics for the remaining items of the PRKOAT have been provided in table 12.

Participant's scores on the PRKOAT ranged from 0 to 64 (mean score 17.11 ± 15.67), with injured athletes (mean score 39.25 ± 14.00) scoring significantly higher (lower knee function) than non-injured athletes (11.93 ± 10.78) ($t_{188} = 12.89$; $p < .01$). Among the participants sampled in this study only 3.9% obtained the maximum score, a score of zero (i.e., ceiling effect), while none of the participants obtained the minimum score (i.e., floor effect). As expected, the scores for the PRKOAT were positively skewed ($Z = 2.45$; $p < .01$). A histogram of participant scores is displayed in figure 18. The skewed distribution is due to the low number of injured participants recruited for this study; however, one of the purposes of this study was to minimize ceiling effects. Therefore, this is not a major limitation.

Table 12. Item Statistics Rasch Partial-Credit Model for Reduced Item Model

Item	Infit	Outfit	b ₀	b ₁	b ₂
1. Pain while straightening knee fully?	0.87	0.66	-4.74	-1.48	-1.21
2. How severe is your knee joint stiffness?	0.98	0.83	-4.49	-3.35	-1.00
3. Pain while lifting a large object off of the floor?	0.88	0.72	-4.14	-1.87	-1.13
4. Pain while jogging at 50% intensity	0.82	0.65	-4.08	-1.96	-1.18
5. Pain with bending knee fully?	1.04	0.80	-3.99	-2.20	-0.93
6. Pain while going up or down stairs?	0.71	0.56	-3.96	-1.62	-0.87
7. Pain while sitting or lying?	1.13	0.86	-3.87	-3.07	-1.30
8. Pain while back pedaling (i.e., running backwards)?	0.66	0.53	-3.87	-2.37	-1.17
9. What is the highest level of activity that you can perform without significant swelling in your knee?	0.92	1.13	-3.80	-2.33	-3.09

10. How often do you experience knee pain?	0.70	0.71	-3.71	-0.82	1.22
11. Pain while bending to the floor and picking up a small object (such as a penny) off the floor?	0.97	0.73	-3.65	-2.16	-1.22
12. Pain while performing standard squats (body weight only)	0.75	0.67	-3.41	-1.80	-0.70
13. Pain while twisting/pivoting on your knee?	0.84	0.71	-3.39	-2.35	-0.34
14. How severe is your knee joint stiffness after sitting, lying, or resting later in the day?	1.05	0.93	-3.31	-3.29	-0.95
15. Pain with jumping forward and landing on your injured knee?	0.68	0.58	-3.23	-1.29	-1.06
16. Compared to your non-injured peers (e.g., teammates or friends of similar athletic ability) how would you rate your usual level of physical activity?	1.41	1.39	-3.18	0.61	n/a
17. Pain while running at maximum speed?	0.85	0.76	-3.07	-1.82	-0.49
18. Pain while standing for an extended period of time (minimum 1 hour)?	1.13	1.00	-3.01	-1.39	-0.13
19. Pain while performing lower-body resistance training exercises with weight (such as lunges, step-ups, or squats)?	0.59	0.53	-2.99	-1.02	-0.50

20. To what extent are you troubled with lack of confidence in your knee when engaging in intense physical activity (such as during practice or games)?	0.75	0.62	-2.99	-2.49	-2.00
21. To what extent do you feel anxious about performing certain activities because of your injured knee?	0.81	0.56	-2.57	-3.20	-1.73
22. Pain while sitting with your knee bent (at any degree) for an extended period of time (minimum 1 hour)?	0.86	0.82	-2.43	-0.73	-0.20
23. Pain while performing single leg squats on your injured leg (body weight only)	0.64	0.51	-2.37	-1.02	-1.13
24. Compared to your healthy knee, how would you rate your injured knee's overall level of function?	1.01	1.00	-2.24	1.89	n/a
25. Compared to your non-injured peers (e.g., teammates or friends of similar athletic ability) how would you rate your overall knee function?	1.03	1.00	-2.23	0.79	n/a
26. Pain with kneeling on your injured knee?	0.88	0.62	-1.88	-1.02	-1.24
27. How often are you aware of your knee problems?	0.92	0.98	-1.84	-1.22	0.47

Note. Item statistics are presented as logit values. b₀ corresponds with the most extreme option (e.g., almost always)

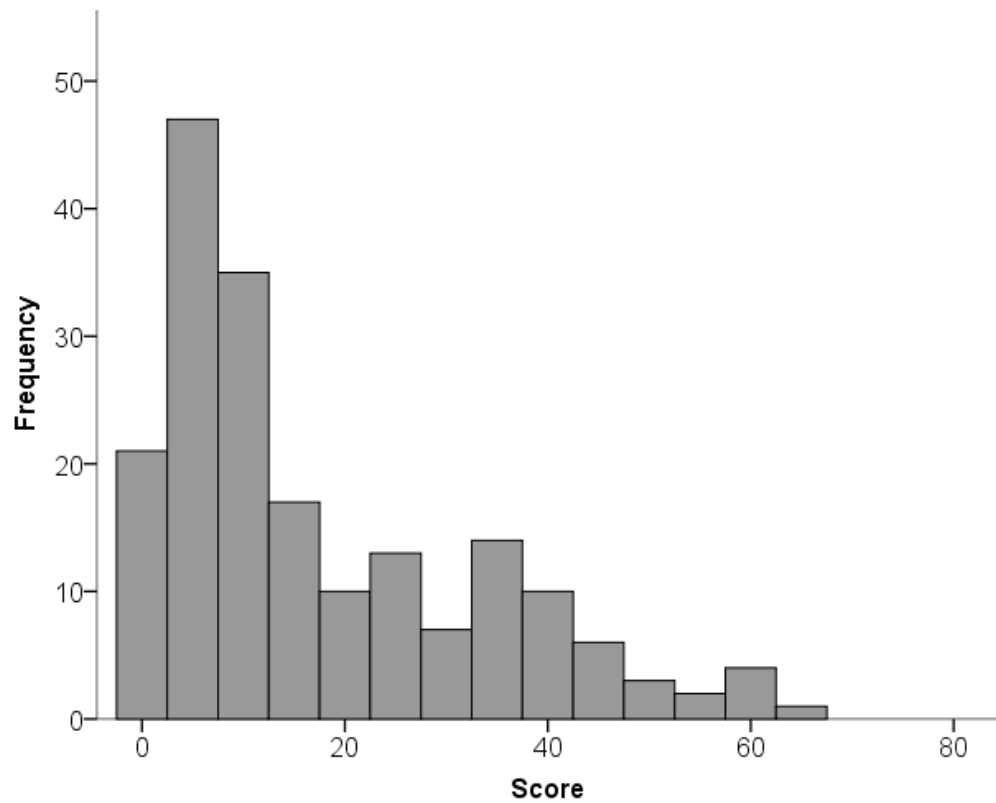


Figure 18. Score Distribution for the Patient Reported Knee Outcomes Assessment Tool (PRKOAT). Scores range from 0 (no problems) to 78 (severe knee dysfunction).

The desired precision for the PRKOAT was set at a conditional standard error of less than 0.30 logits. Conditional standard error of measurement is calculated as the inverse of the test information function. Precision of the PRKOAT is acceptable when function ability is low, however, as θ increases

beyond -0.25 logits the precision of the PRKOAT begins to decrease (see figure 19). The measurement precision for the PRKOAT was within acceptable limits for approximately 39% of the sample. It should also be noted, however, that 67% of the sample had a conditional standard error less than 0.40 logits.

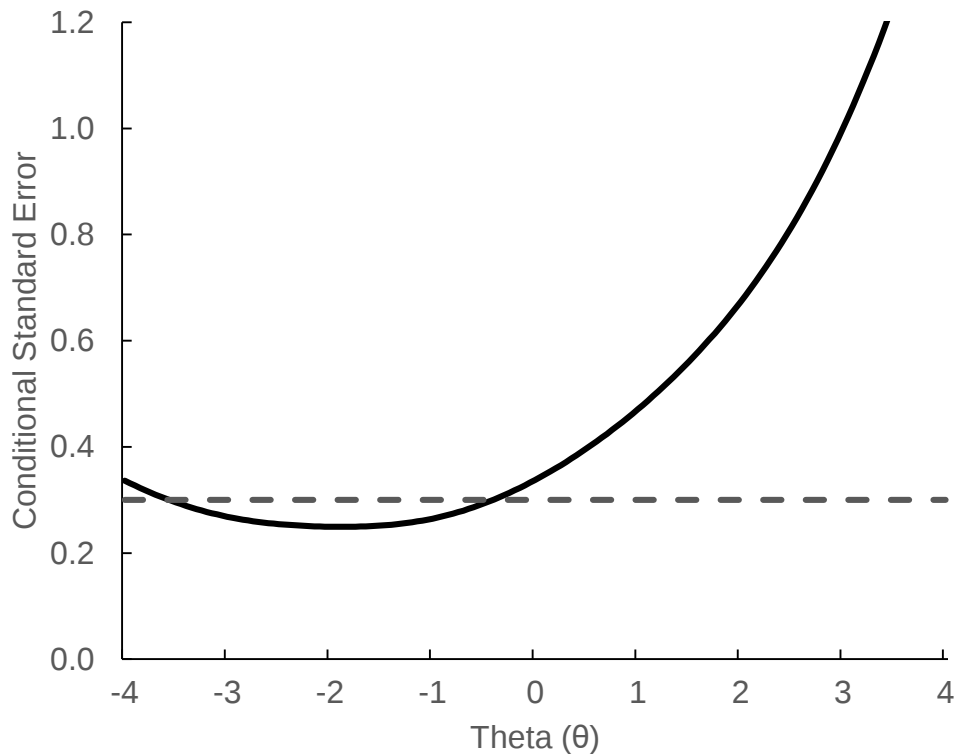


Figure 19. Conditional Standard Error of Measurement for the PRKOAT. The dashed grey line is positioned at 0.3 which indicates acceptable levels of measurement error. The solid black line indicates the degree of measurement error at a particular level of theta (θ).

Discussion

Overall, the results from this study indicate that the PRKOAT can be a useful instrument for measuring an athlete's knee-related function following injury. Known-group difference validity evidence combined with the high Cronbach's α suggests the instrument is suitable for evaluating changes in function. More studies, however, are needed to confirm these results and provide additional sources of validity evidence. Furthermore, responsiveness of this instrument should be assessed both short and long term. Compared to the IKDC-SKF and the KOOS, the PRKOAT had fewer ceiling effects, which should provide enhanced evaluation of knee function throughout the rehabilitation process.

The final revised version of the PRKOAT contained 27 items with acceptable model-data fit. Of the 27 items contained in the PRKOAT, only 4 had disordered category response options: [1] Pain while performing single leg squats on your injured leg (body weight only), [2] Pain with kneeling on your injured knee [3] To what extent do you feel anxious about performing certain activities because of your injured knee, and [4] What is the highest level of activity that you can perform without significant swelling in your knee. The remaining 23 items had relatively good category function. The disordered categories for each of the 4 item with category dysfunction all occurred between the moderate-severe and the mild-moderate boundaries.

One possible explanation is that the number of category options selected for these items may not be appropriate. Research investigating the number of optimal categories for rating scale questions provides conflicting evidence about

the optimal number of response options. While some studies suggest that increasing the number of category options can increase the reliability of an instrument, (E. R. Masters, 1974) other studies have reported the opposite findings (Matell & Jacoby, 1971). A second potential explanation is related to the target sample of the study. Majority of the participants included in this study were generally healthy with no current injuries ($n = 167$; 82.3%) resulting in few participants responding in the most severe category. Because of this, estimation of boundary locations for the extreme categories may be biased. In the clinical setting this instrument is intended to be used on athletes who have recently been injured or are currently recovering from injury and initially will be expected to have greater functional limitations, making better use of the extreme categories.

Two major advantages of this study were the use of advanced measurement theories (i.e., Rasch measurement theory) and the target sample (i.e., athletes). Many of the current PRO measures were not originally developed for athletic populations (e.g., KOOS); those that were developed for athletes (e.g., IKDC-SKF) were not calibrated using athletes, rather the general population was used. As stated previously, studies have demonstrated key differences between athletes and non-athletes on measures of psychological and physiological well-being, key components of measuring patient function. Despite having strong content validity, and the use of Rasch modeling, when the IKDC-SKF was testing in athletes, the items yielded poor model-data fit (Farnsworth II et al., n.d.). These factors may be a contributing factor to the abundant ceiling effects identified in both the KOOS and IKDC-SKF (Fries et al., 2011; Roos &

Toksvig-Larsen, 2003). Ceiling effects as high as 28% have been identified in the KOOS with ceiling effects as high as 15% for the IKDC-SKF. Minimizing ceiling effects represents a critical issue for PRO measures considering their repeated use for tracking changes in function over time. In the present study, only 3.9% of participants achieved the maximum score, a substantial improvement. It should also be noted that majority of the participants in this study were currently in-season with minimal or no injuries. Thus, with injured athletes, whom this measure is targeted for, the frequency of ceiling effects may even be lower.

An alternative solution for addressing ceiling effects, often used, is functional-based-field testing such as one-leg hop for distance test, timed single-leg hop test, y-balance test, and percentage leg press. These functional-based tests however, have poor correlations with PRO measures indicating the two assessment strategies are measuring different latent constructs (Burton, 2017). Another limitation of functional-based-field tests is the inability to assess psychosocial dimensions that are considered key components of patient function and overall well-being as defined by the World Health Organization (World Health Organization, 2001). While these functional-based-field tests may provide useful supplemental information, they should not be considered an alternative to traditional PRO measures.

Despite the numerous benefits associated with inclusion of PRO measures in clinical practice athletic trainers often report that administration time is a common barrier to implementation. Compared to the KOOS which contains 42 items, the PRKOAT only contains a total of 27 items, a 35% decrease in the

total number of items; which should lead to decreased administration time. Furthermore, the PRKOAT was developed and hosted online using Google Forms. Google Forms is a free software program that allows for development of survey instruments similar to Survey Monkey, which can be accessed from not only computers, but also tablets and cellular devices. Because the PRKOAT is hosted online, patients can complete the PRO measure before coming into the clinical setting, allowing time for the clinician to review the patient's information and enhance communication between the patient and clinician. This has the potential to decrease clinician burden making the PRKOAT an attractive option for clinicians.

Another commonly reported barrier for implementation of PRO measures in clinical practice is that many of the items included on this instruments are not relevant to athletes. For example, on the KOOS activities of daily living subscale many items such as difficulty with getting in/out of a car, going shopping, getting in/out of bath, heavy domestic duties, and light domestic duties do not apply to athlete populations. In many cases athletes often live on campus in dorm rooms where these questions may not apply. Many of the items added to the PRKOAT were more sport specific and may help to enhance the clinician's ability to evaluate patients. Single leg squats for instance are often included in rehabilitation programs following ACL injury (Tagesson, Öberg, Good, & Kvist, 2008) to help improve knee stability and lower extremity strength in the involved limb. The inclusion of this particular item, as well as the addition of sport-specific items allows for clinicians to assess the patient's comfort and ability to perform

these skills, which can then be documented over time to ensure adequate treatment outcomes. Additional changes which occurred during the pilot phase of this study to the quality of life items to be more specific to patient's needs should also serve to maximize the relevancy of the items making this instrument very valuable for both clinicians and student-athletes as they recover from injury.

Limitations

This study was separated into two distinct phases, the development/pilot-test phase and the validation phase. Modifications to the original instrument were based upon information obtained during a small pilot study, where sample size was understandably small. As a result, inferences from the pilot study may be biased; however, many of the changes to the original instrument were supported by data collected in the validation phase from a much larger sample. In addition, while there was some interaction and discussion with participants regarding the study instrument during the pilot test, the data for this study relied heavily on self-report information collected anonymously from participants online. Due to the lack of physical contact with these participants there was no way to verify the validity of the participants' injury claims. Lastly, although every effort was made to collect data on a diverse sample, majority of the sample population in this study was division III college athletes. Therefore, the information may not be generalizable to other groups of athletes. More studies should be conducted to investigate the quality of this instrument in those groups.

Conclusion

The PRKOAT is a newly developed PRO measure for assessing knee function in athletes. A major advantage of this study was the use of advanced measurement theory (i.e., Rasch modeling) and the targeted population. Compared to commonly used PRO measures (e.g., KOOS, IKDC-SKF) the PRKOAT has greatly reduced ceiling effects in athletic populations. Clinicians should consider using the PRKOAT to evaluate PRO throughout the rehabilitation process following knee-related injuries. Future studies are needed to investigate the reliability and responsiveness of this instrument to various knee injuries.

REFERENCES

- Aglietti, P., Giron, F., Buzzi, R., Biddau, F., & Sasso, F. (2004). Anterior cruciate ligament reconstruction: bone-patellar tendon-bone compared with double semitendinosus and gracilis tendon grafts: a prospective, randomized clinical trial. *The Journal of Bone and Joint Surgery*, 86(10), 2143-2155.
- Burt, C. W., & Overpeck, M. D. (2001). Emergency visits for sports-related injuries. *Annals of Emergency Medicine*, 37(3), 301-308.
- Burton, C. (2017). *Can the IKDC Subjective Patient Outcome Predict Limb Asymmetries in Clinician Reported Outcomes?* Paper presented at the National Athletic Trainers' Association 68th Clinical Symposia & AT Expo, Houston, TX. Free Communication Oral Presentations.
- Collins, N. J., Prinsen, C. A. C., Christensen, R., Bartels, E. M., Terwee, C. B., & Roos, E. M. (2016). Knee Injury and Osteoarthritis Outcome Score (KOOS): systematic review and meta-analysis of measurement properties. *Osteoarthritis and Cartilage*, 24(8), 1317-1329.
- Farnsworth II, J., Evans, T. A., Binkley, H., & Kang, M. (n.d.). *Psychometric evaluation of patient reported knee outcomes for assessment of knee-related function*. Dissertation Data. Health and Human Performance. Middle Tennessee State University.
- Flosadottir, V., Frobell, R., Roos, E. M., & Ageberg, E. (2017). Impact of treatment strategy and physical performance on knee self-efficacy in individuals with anterior cruciate ligament injury. *Osteoarthritis and Cartilage*, 25, S394.

- Fries, J., Rose, M., & Krishnan, E. (2011). The PROMIS of better outcome assessment: responsiveness, floor and ceiling effects, and Internet administration. *The Journal of Rheumatology*, 38(8), 1759-1764.
- Gage, B. E., McIlvain, N. M., Collins, C. L., Fields, S. K., & Comstock, R. D. (2012). Epidemiology of 6.6 million knee injuries presenting to United States emergency departments from 1999 through 2008. *Academic Emergency Medicine*, 19(4), 378-385.
- Gobbi, A., & Francisco, R. (2006). Factors affecting return to sports after anterior cruciate ligament reconstruction with patellar tendon and hamstring graft: a prospective clinical investigation. *Knee Surgery, Sports Traumatology, Arthroscopy*, 14(10), 1021-1028.
- Institute of Medicine (US) Committee on Quality of Health Care in America. (2001). *Crossing the quality chasm: A new health system for the 21st century* (Vol. 323). Washington DC: National Academic Press.
- Irrgang, J. J., Anderson, A. F., Boland, A. L., Harner, C. D., Kurosaka, M., Neyret, P., . . . Shelborne, K. D. (2001). Development and validation of the international knee documentation committee subjective knee form. *The American Journal of Sports Medicine*, 29(5), 600-613.
- Johanson, N. A., Liang, M. H., Daltroy, L., Rudicel, S., & Richmond, J. (2004). American Academy of Orthopaedic Surgeons lower limb outcomes assessment instruments: reliability, validity, and sensitivity to change. *The Journal of Bone and Joint Surgery*, 86(5), 902-909.

Kincaid, J. P., Fishburne Jr, R. P., Rogers, R. L., & Chissom, B. S. (1975).

Derivation of new readability formulas (automated readability index, fog count and flesch reading ease formula) for navy enlisted personnel.

Kvist, J., Ek, A., Sporrstedt, K., & Good, L. (2005). Fear of re-injury: a hindrance for returning to sports after anterior cruciate ligament reconstruction. *Knee Surgery, Sports Traumatology, Arthroscopy*, 13(5), 393-397.

Lee, D. Y., Karim, S. A., & Chang, H. C. (2008). Return to sports after anterior cruciate ligament reconstruction -- a review of patients with minimum 5-year follow-up. *Annals Academy of Medicine Singapore*, 37(4), 273-278.

Linacre, J. M. (2002). Optimizing rating scale category effectiveness. *Journal of Applied Measurement*, 3(1), 85-106.

Masters, E. R. (1974). The relationship between the number of response categories and reliability of Likert-type questionnaires. *Journal of Educational Measurement*, 11(1), 49-53.

Matell, M. S., & Jacoby, J. (1971). Is there an optimal number of alternatives for Likert scale items? Study I: Reliability and validity. *Educational and Psychological Measurement*, 31(3), 657-674.

Paterno, M. V., Ford, K. R., Myer, G. D., Heyl, R., & Hewett, T. E. (2007). Limb asymmetries in landing and jumping 2 years following anterior cruciate ligament reconstruction. *Clinical Journal of Sport Medicine*, 17(4), 258-262.

Roos, E. M., Roos, H. P., Lohmander, L. S., Ekdahl, C., & Beynnon, B. D. (1998). Knee Injury and Osteoarthritis Outcome Score (KOOS)—development of a

self-administered outcome measure. *Journal of Orthopaedic & Sports Physical Therapy*, 28(2), 88-96.

Roos, E. M., & Toksvig-Larsen, S. (2003). Knee injury and Osteoarthritis Outcome Score (KOOS)-validation and comparison to the WOMAC in total knee replacement. *Health and Quality of Life Outcomes*, 1(1), 17.

Snyder, A. R., Martinez, J. C., Bay, R. C., Parsons, J. T., Sauers, E. L., & McLeod, T. C. V. (2010). Health-related quality of life differs between adolescent athletes and adolescent nonathletes. *Journal of Sport Rehabilitation*, 19(3), 237-248.

Tagesson, S., Öberg, B., Good, L., & Kvist, J. (2008). A comprehensive rehabilitation program with quadriceps strengthening in closed versus open kinetic chain exercise in patients with anterior cruciate ligament deficiency: a randomized clinical trial evaluating dynamic tibial translation and muscle function. *The American Journal of Sports Medicine*, 36(2), 298-307.

Wiger, P., Brandsson, S., Kartus, J., Eriksson, B., & Karlsson, J. (1999). A comparison of results after arthroscopic anterior cruciate ligament reconstruction in female and male competitive athletes A two-to five-year follow-up of 429 patients. *Scandinavian Journal of Medicine & Science in Sports*, 9(5), 290-295.

World Health Organization. (2001). *International Classification of Functioning, Disability and Health*. Geneva: World Health Organization.

- Wright, B. D., Linacre, J. M., Gustafson, J., & Martin-Lof, P. (1994). Reasonable mean-square fit values. *Rasch Measurement Transactions*, 8(3), 370.
- Yamazaki, J., Muneta, T., Ju, Y. J., & Sekiya, I. (2009). Differences in kinematics of single leg squatting between anterior cruciate ligament-injured patients and healthy controls. *Knee Surgery, Sports Traumatology, Arthroscopy*, 18(1), 56-63.
- Young, K., White, P., & McTeer, W. (1994). Body talk: Male athletes reflect on sport, injury, and pain. *Sociology of Sport Journal*, 11(2), 175-194.

CHAPTER V

OVERALL CONCLUSION

Patient-reported outcomes (PRO) measures are an integral component of the injury evaluation process. In the athletic training setting, the top reported benefits of PRO measures by athletic trainers are the enhancement of communication with patients and other health care professionals, assistance with directing patient care, and increased examination efficiency (Valier, Jennings, Parsons, & Vela, 2014). Although there are numerous advantages to implementation of these instruments some report that current instruments are sometimes confusing and the items are not relevant to their patients (Valier et al., 2014). Among these reported barriers, studies have demonstrated that many of the current instruments are not appropriate for athletes due to design limitations (i.e., sample population and ceiling effects) (Fries et al., 2011; Roos & Toksvig-Larsen, 2003).

In an effort to address these barriers and limitations a new instrument capable of measuring function in athletes was needed. Therefore, the overall purpose of this project was to evaluate the limitations of current instruments and to develop a new and improved instrument that was able to be used for assessing athlete's function in the clinical setting. To accomplish this goal two studies were devised.

The first study involved examination of commonly used knee-related PRO measures which included: The Knee Injury and Osteoarthritis Outcomes Score (KOOS), the International Knee Documentation Committee's Subjective Knee

Form (IKDC-SKF), and the Marx Activity Rating Scale using advanced statistical models to determine the measurement properties of each instrument. The results from the first study indicated that a lack of sufficiently challenging items was likely the cause of ceiling effects in these instruments. Although functional-based testing is often used as a Band-Aid solution, these tests have poor correlations with PRO measures (Burton, 2017) and fail to assess sociological and psychological variables related to patient well-being which are important components of function (World Health Organization, 2001). Moreover, athletes often report higher quality of life compared to the general population which exacerbates the problem of high ceiling effects (Lam, K. C., Valier A. R., Bay, C. R., & McLeod, T. C. V., 2013; Snyder et al., 2010). To address this concern new items were developed to minimize ceiling effects and provide a more precise estimate of function for athletes.

In the second study, a new instrument, the Patient Reported Knee Outcomes Assessment Tool (PRKOAT) was developed by combining items that had acceptable model-data fit from the KOOS and IKDC-SKF (no items from the Marx activity rating scale were retained due to poor measurement properties) with the newly developed items. The PRKOAT was pilot tested in a small sample of in-season athletes where participants were asked to provide feedback about the overall quality and usefulness of the instrument. Using feedback from the athletes and results from the pilot test the instrument was refined and re-evaluated using a larger sample.

Results from the second study indicated that compared to the KOOS and IKDC-SKF there are fewer participants obtaining the maximum score. This has significant and immediate implications for clinical practice. Furthermore, the item step difficulty of the most challenging item in the PRKOAT exceeds those of previous instruments by 1.32 logits. This suggests that the PRKOAT is capable of measuring higher levels of function than previous instruments.

While the new PRKOAT has good potential as a functional outcome measure there are still concerns regarding application of these instruments in clinical practice. Currently, many clinicians either only use PRO measures during the initial injury stages, or they don't use them at all (Valier et al., 2014). This can make it difficult to objectively evaluate the effectiveness of rehabilitation programs because each clinician is using different outcome measures limiting comparisons. While the inclusion of more difficult and challenging items improved the quality of PRO measures, there are still concerns regarding the precision of the PRKOAT with patients who have very high levels of function. Conditional standard errors of measurement indicate that as function levels improve, similar to other instruments, the error rate increases. Because of the use of Rasch modeling items can be added at a later date, without diminishing the measurement properties of the current instrument due to local item independence.

In summary, the PRKOAT, which was developed using sound measurement principles is a useful measure of athlete knee function. Future studies are needed to assess the reliability of this instrument both short and long

term. In addition, additional validity evidence is needed to assess the sensitivity and responsiveness of the instrument to changes in knee function over time.

Current evidence suggest that the instrument is capable of detecting changes in functional ability between individuals with differing levels of function.

REFERENCES

- Ader, D. N. (2007). Developing the Patient-Reported Outcomes Measurement Information System (PROMIS). *Medical Care*, 45(5), Suppl 1, S1-S2.
- Agel, J., Dick, R., Nelson, B., Marshall, S. W., & Dompier, T. P. (2007). Descriptive epidemiology of collegiate women's ice hockey injuries: National Collegiate Athletic Association Injury Surveillance System, 2000-2001 through 2003-2004. *Journal of Athletic Training*, 42(2), 249-254.
- Agel, J., Dompier, T. P., Dick, R., & Marshall, S. W. (2007). Descriptive epidemiology of collegiate men's ice hockey injuries: National Collegiate Athletic Association Injury Surveillance System, 1988-1989 through 2003-2004. *Journal of Athletic Training*, 42(2), 241-248.
- Agel, J., Evans, T. A., Dick, R., Putukian, M., & Marshall, S. W. (2007). Descriptive epidemiology of collegiate men's soccer injuries: National Collegiate Athletic Association Injury Surveillance System, 1988-1989 through 2002-2003. *Journal of Athletic Training*, 42(2), 270-277.
- Agel, J., Olson, D. E., Dick, R., Arendt, E. A., Marshall, S. W., & Sikka, R. S. (2007). Descriptive epidemiology of collegiate women's basketball injuries: National Collegiate Athletic Association Injury Surveillance System, 1988-1989 through 2003-2004. *Journal of Athletic Training*, 42(2), 202-210.
- Agel, J., Palmieri-Smith, R. M., Dick, R., Wojtys, E. M., & Marshall, S. W. (2007). Descriptive epidemiology of collegiate women's volleyball injuries: National Collegiate Athletic Association Injury Surveillance System, 1988-1989 through 2003-2004. *Journal of Athletic Training*, 42(2), 295-302.

- Agel, J., Ransone, J., Dick, R., Oppliger, R., & Marshall, S. W. (2007). Descriptive epidemiology of collegiate men's wrestling injuries: National Collegiate Athletic Association Injury Surveillance System, 1988-1989 through 2003-2004. *Journal of Athletic Training*, 42(2), 303-310.
- Aglietti, P., Giron, F., Buzzi, R., Biddau, F., & Sasso, F. (2004). Anterior cruciate ligament reconstruction: bone-patellar tendon-bone compared with double semitendinosus and gracilis tendon grafts: a prospective, randomized clinical trial. *The Journal of Bone and Joint Surgery*, 86(10), 2143-2155.
- American Academy of Orthopaedic Surgeons. (2016). Patient Reported Outcome Measures. Retrieved from http://www.aaos.org/Quality/Performance_Measures/Patient_Reported_Outcome_Measures/
- Andrich, D. (1978). A rating formulation for ordered response categories. *Psychometrika*, 43(4), 561-573.
- Andrich, D. (2004). Controversy and the Rasch model: a characteristic of incompatible paradigms? *Medical Care*, 42(1 suppl), 7-16.
- Annas, G. J. (1975). *The rights of hospital patients: American Civil Liberties Union handbook*. New York: Avon.
- Ansley, T. N., & Forsyth, R. A. (1985). An examination of the characteristics of unidimensional IRT parameter estimates derived from two-dimensional data. *Applied Psychological Measurement*, 9(1), 37-48.
- Bach, B. R., & Boonos, C. L. (2001). Anterior cruciate ligament reconstruction. *Association of periOperative Registered Nurses Journal*, 74(2), 151-164.

- Badley, E. M., Thompson, R. P., & Wood, P. H. (1978). The prevalence and severity of major disabling conditions-a reappraisal of the Government Social Survey on the handicapped and impaired in Great Britain. *International Journal of Epidemiology*, 7(2), 145-151.
- Barber, S. D., & Noyes, F. R. (March, 16th, 2015). The Cincinnati Knee Rating System. *Orthopaedics*. Retrieved from <http://clinicalgate.com/the-cincinnati-knee-rating-system/#bb0295>
- Barber, S. D., Noyes, F. R., Mangine, R. E., & Hartman, W. (1990). Quantitative assessment of functional limitations in normal and anterior cruciate ligament-deficient knees. *Clinical Orthopaedics and Related Research*, 255, 204-214.
- Bellamy, N. (1988). Validation study of WOMAC: a health status instrument for measuring clinically-important patient-relevant outcomes following total hip or knee arthroplasty in osteoarthritis. *Journal of Orthopedics & Rheumatology*, 1, 95-108.
- Binkley, J. (1999). Measurement of functional status, progress and outcome in orthopaedic clinical practice. *Journal of Orthopaedic & Sports Physical Therapy*, 11(1), 14-21.
- Birnbaum, A. (1957). *Efficient design and use of tests of a mental ability for various decision-making problems* (7755-23). Retrieved from Randolph Air Force Base, Texas: Air University, School of Aviation Medicine:

- Birnbaum, A. (1958a). *Further considerations of efficiency in tests of a mental ability* (7755-23). Retrieved from USAF School of Aviation Medicine, Randolph Air Force Base, Texas
- Birnbaum, A. (1958b). *On the estimation of mental ability* (7755-23). Retrieved from USAF School of Aviation Medicine, Randolph Air Force Base, Texas
- Birnbaum, A. (1968). Some latent trait models and their use in inferring an examinee's ability. In F. Lord & M. Novick (Eds.), *Statistical theories of mental test scores* (pp. 397-479). Reading, MA: Addison-Wesley.
- Bock, R. D., & Aitkin, M. (1981). Marginal maximum likelihood estimation of item parameters: Application of an EM algorithm. *Psychometrika*, 46(4), 443-459.
- Bombardier, C., Melfi, C. A., Paul, J., Green, R., Hawker, G., Wright, J., & Coyte, P. (1995). Comparison of a generic and a disease-specific measure of pain and physical function after knee replacement surgery. *Medical Care*, AS131-AS144.
- Briggs, K. K., Lysholm, J., Tegner, Y., Rodkey, W. G., Kocher, M. S., & Steadman, J. R. (2009). The reliability, validity, and responsiveness of the Lysholm score and Tegner activity scale for anterior cruciate ligament injuries of the knee 25 years later. *The American Journal of Sports Medicine*, 37(5), 890-897.
- Burt, C. W., & Overpeck, M. D. (2001). Emergency visits for sports-related injuries. *Annals of Emergency Medicine*, 37(3), 301-308.

- Burton, C. (2017). *Can the IKDC Subjective Patient Outcome Predict Limb Asymmetries in Clinician Reported Outcomes?* Paper presented at the National Athletic Trainers' Association 68th Clinical Symposia & AT Expo, Houston, TX.
- Busfield, B. T., Kharrazi, F. D., Starkey, C., Lombardo, S. J., & Seegmiller, J. (2009). Performance outcomes of anterior cruciate ligament reconstruction in the National Basketball Association. *The Journal of Arthroscopic & Related Surgery*, 25(8), 825-830.
- Cella, D., Yount, S., Rothrock, N., Gershon, R., Cook, K., Reeve, B., . . . Rose, M. (2007). The Patient-Reported Outcomes Measurement Information System (PROMIS): Progress of an NIH roadmap cooperative group during its first two years. *Medical Care*, 45(5), S3-S11.
- Chan, R. (2005). How severe should symptoms be before someone is said to be suffering from post-concussion syndrome? An exploratory study with self-reported checklist using Rasch analysis. *Brain Injury*, 19(13), 1117-1124.
- Coghill, R. C., McHaffie, J. G., & Yen, Y.-F. (2003). Neural correlates of interindividual differences in the subjective experience of pain. *Proceedings of the National Academy of Sciences*, 100(14), 8538-8542.
- Coker, C. A. (2013). *Motor learning & control for practitioners with online labs* (3rd ed.). Scottsdale Arizona: Holcomb Hathaway, Publishers.
- Collins, N., Prinsen, C., Christensen, R., Bartels, E., Terwee, C., & Roos, E. (2016). Knee Injury and Osteoarthritis Outcome Score (KOOS):

systematic review and meta-analysis of measurement properties.

Osteoarthritis and Cartilage, 24(8), 1317-1329.

Conway, J. M., & Huffcutt, A. I. (2003). A review and evaluation of exploratory factor analysis practices in organizational research. *Organizational Research Methods*, 6(2), 147-168.

Coute, R. A., Ehrenfeld, J. M., Gupta, D. K., Terekhov, M. A., & Wanderer, J. P. (2017). Electronically self-assessed functional capacity and exercise testing: A comparison of the Duke Activity Status Index and Patient-Reported Outcomes Measurement Information System tools. *American Heart Journal*, 188, 82-86.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297-332.

Curb, J. D., Ceria-Ulep, C. D., Rodriguez, B. L., Grove, J., Guralnik, J., Willcox, B. J., . . . Chen, R. (2006). Performance-based measures of physical function for high-function populations. *Journal of the American Geriatrics Society*, 54(5), 737-742.

Daniel, D. M., Stone, M. L., Dobson, B. E., Fithian, D. C., Rossman, D. J., & Kaufman, K. R. (1994). Fate of the ACL-injured patient a prospective outcome study. *The American Journal of Sports Medicine*, 22(5), 632-644.

Deroche, T., Woodman, T., Stephan, Y., Brewer, B. W., & Le Scanff, C. (2011). Athletes' inclination to play through pain: a coping perspective. *Anxiety, Stress, & Coping*, 24(5), 579-587.

- Deyo, R. A., & Patrick, D. L. (1989). Barriers to the use of health status measures in clinical investigation, patient care, and policy research. *Medical Care*, S254-S268.
- Dick, R., Ferrara, M. S., Agel, J., Courson, R., Marshall, S. W., Hanley, M. J., & Reifsteck, F. (2007). Descriptive epidemiology of collegiate men's football injuries: National Collegiate Athletic Association Injury Surveillance System, 1988–1989 through 2003–2004. *Journal of Athletic Training*, 42(2), 221-233.
- Dick, R., Hertel, J., Agel, J., Grossman, J., & Marshall, S. W. (2007). Descriptive epidemiology of collegiate men's basketball injuries: National Collegiate Athletic Association Injury Surveillance System, 1988-1989 through 2003-2004. *Journal of Athletic Training*, 42(2), 194-201.
- Dick, R., Lincoln, A. E., Agel, J., Carter, E. A., Marshall, S. W., & Hinton, R. Y. (2007). Descriptive epidemiology of collegiate women's lacrosse injuries: National Collegiate Athletic Association Injury Surveillance System, 1988–1989 through 2003–2004. *Journal of Athletic Training*, 42(2), 262-269.
- Dick, R., Putukian, M., Agel, J., Evans, T. A., & Marshall, S. W. (2007). Descriptive epidemiology of collegiate women's soccer injuries: National Collegiate Athletic Association Injury Surveillance System, 1988-1989 through 2002-2003. *Journal of Athletic Training*, 42(2), 278-285.
- Dick, R., Romani, W. A., Agel, J., Case, J. G., & Marshall, S. W. (2007). Descriptive epidemiology of collegiate men's lacrosse injuries: National

- Collegiate Athletic Association Injury Surveillance System, 1988-1989 through 2003-2004. *Journal of Athletic Training*, 42(2), 255-261.
- Dick, R., Sauers, E. L., Agel, J., & Keuter, G. (2007). Descriptive epidemiology of collegiate men's baseball injuries: National Collegiate Athletic Association Injury Surveillance System, 1988-1989 through 2003-2004. *Journal of Athletic Training*, 42(2), 183-193.
- Dorans, N. J., & Kingston, N. M. (1985). The effects of violations of unidimensionality on the estimation of item and ability parameters and on item response theory equating of the GRE verbal scale. *Journal of Educational Measurement*, 22(4), 249-262.
- Douglas, J., Price, M., & Peters, D. (2012). A systematic review of physical fitness, physiological demands and biomechanical performance in equestrian athletes. *Comparative Exercise Physiology*, 8(1), 53-62.
- Drasgow, F., & Parsons, C. K. (1983). Application of unidimensional item response theory models to multidimensional data. *Applied Psychological Measurement*, 7(2), 189-199.
- Embretson, S. E., & Reise, S. P. (2013). *Item response theory for psychologists*. Mahwah, New Jersey: Psychology Press.
- Engström, B., Forssblad, M., Johansson, C., & Tornkvist, H. (1990). Does a major knee injury definitely sideline an elite soccer player? *The American Journal of Sports Medicine*, 18(1), 101-105.

Espallargues, M., Valderas, J. M., & Alonso, J. (2000). Provision of feedback on perceived health status to health care professionals: a systematic review of its impact. *Medical Care*, 38(2), 175-186.

Farnsworth II, J., Evans, T. A., Binkley, H., & Kang, M. (n.d.). *Psychometric evaluation of patient reported knee outcomes for assessment of knee-related function*. Dissertation Data. Health and Human Performance. Middle Tennessee State University.

Farnsworth II, J. Dargo, L., Ragan, B. G., Kang, M. (2017). Reliability of computerized neurocognitive tests for concussion assessment: A meta-analysis. *Journal of Athletic Training*, 52(6), 826-833. doi: 10.4085/1062-6050-52.6.03

Farnsworth II, J., McElhiney, D., David, S., Sinha, G., & Ragan, B. G. (2013). Objective functional assessment after a head injury using movement and activity in physical space scores: a case report. *Journal of Athletic Training*, 49(4), 568-575.

Fennessy, L. M. (1995). *The impact of local dependencies on various IRT outcomes*. (Ph.D. Dissertation), University of Massachusetts at Amherst, Ann Arbor. Retrieved from <http://search.proquest.com/docview/304217630> (9524701)

Flosadottir, V., Frobell, R., Roos, E. M., & Ageberg, E. (2017). Impact of treatment strategy and physical performance on knee self-efficacy in individuals with anterior cruciate ligament injury. *Osteoarthritis and Cartilage*, 25, S394. doi:10.1016/j.joca.2017.02.678

- Fries, J., Rose, M., & Krishnan, E. (2011). The PROMIS of better outcome assessment: responsiveness, floor and ceiling effects, and Internet administration. *The Journal of Rheumatology*, 38(8), 1759-1764.
- Fulkerson, J. P. (2004). Nonarthritic anterior knee pain. In *Disorders of the patellafemoral joint* (4th ed., pp. 143-155). Philadelphia: Lippincott Williams & Wilkins.
- Gage, B. E., McIlvain, N. M., Collins, C. L., Fields, S. K., & Comstock, R. D. (2012). Epidemiology of 6.6 million knee injuries presenting to United States emergency departments from 1999 through 2008. *Academic Emergency Medicine*, 19(4), 378-385. doi:10.1111/j.1553-2712.2012.01315.x
- Gobbi, A., & Francisco, R. (2006). Factors affecting return to sports after anterior cruciate ligament reconstruction with patellar tendon and hamstring graft: a prospective clinical investigation. *Knee Surgery, Sports Traumatology, Arthroscopy*, 14(10), 1021-1028.
- Greenfield, S., Kaplan, S., & Ware, J. J. E. (1985). Expanding patient involvement in care effects on patient outcomes. *Annals of Internal Medicine*, 102(4), 520-528. doi:10.7326/0003-4819-102-4-520
- Greenhalgh, J., & Meadows, K. (1999). The effectiveness of the use of patient-based measures of health in routine practice in improving the process and outcomes of patient care: a literature review. *Journal of Evaluation in Clinical Practice*, 5(4), 401-416.

- Guccione, A. A. (1991). Physical therapy diagnosis and the relationship between impairments and function. *Physical Therapy*, 71(7), 499-503.
- Hambleton, R. K., Swaminathan, H., & Rogers, J. (1991). *Fundamentals of Item Response Theory*. Newbury Park, California: SAGE Publications, Inc.
- Han, H. S., Seong, S. C., Lee, S., & Lee, M. C. (2008). Anterior cruciate ligament reconstruction. *Clinical Orthopaedics and Related Research*, 466(1), 198-204.
- Harris, A. I. (1971). *Handicapped and impaired in Great Britain*. London: Highline Medical Services Organization.
- Hattie, J. (1985). Methodology review: assessing unidimensionality of tests and items. *Applied Psychological Measurement*, 9(2), 139-164.
- Hefti, E., Müller, W., Jakob, R., & Stäubli, H. (1993). Evaluation of knee ligament injuries with the IKDC form. *Knee Surgery, Sports Traumatology, Arthroscopy*, 1(3-4), 226-234.
- Hefti, F., Gächter, A., Jenny, H., & Morscher, E. (1982). Replacement of the anterior cruciate ligament. *Archives of Orthopaedic and Traumatic Surgery*, 100(2), 83-94.
- Herrmann, S. D., Snook, E. M., Kang, M., Scott, C. B., Mack, M. G., Dompier, T. P., & Ragan, B. G. (2011). Development and validation of a movement and activity in physical space score as a functional outcome measure. *Archives of Physical Medicine and Rehabilitation*, 92(10), 1652-1658.
- Hunter, D. J., & Felson, D. T. (2006). Osteoarthritis. *British Medical Journal*, 332(7542), 639-642.

- Institute of Medicine (US) Committee on Quality of Health Care in America. (2001). *Crossing the quality chasm: A new health system for the 21st century* (Vol. 323). Washington DC: National Academic Press.
- Irrgang, J. J., Anderson, A. F., Boland, A. L., Harner, C. D., Kurosaka, M., Neyret, P., . . . Shelborne, K. D. (2001). Development and validation of the international knee documentation committee subjective knee form. *The American Journal of Sports Medicine*, 29(5), 600-613.
- Jette, A. M. (1994). Physical disablement concepts for physical therapy research and practice. *Physical Therapy*, 74(5), 380-386.
- Johanson, N. A., Liang, M. H., Daltroy, L., Rudicel, S., & Richmond, J. (2004). American Academy of Orthopaedic Surgeons lower limb outcomes assessment instruments: reliability, validity, and sensitivity to change. *The Journal of Bone and Joint Surgery*, 86(5), 902-909.
- Kantz, M. E., Harris, W. J., Levitsky, K., Ware Jr, J. E., & Davies, A. R. (1992). Methods for assessing condition-specific and generic functional status outcomes after total knee replacement. *Medical Care*, 30(5), MS240-MS252.
- Kincaid, J. P., Fishburne Jr, R. P., Rogers, R. L., & Chissom, B. S. (1975). *Derivation of new readability formulas (automated readability index, fog count and flesch reading ease formula) for navy enlisted personnel* (No. RBR-8-75). Naval Technical Training Command Millington TN Research Branch.

- Kreiner, S., & Christensen, K. B. (2007). Validity and objectivity in health-related scales: analysis by graphical loglinear Rasch models. In *Multivariate and mixture distribution Rasch models* (pp. 329-346). New York: Springer.
- Kurosaka, M., Yoshiya, S., & Andrish, J. T. (1987). A biomechanical comparison of different surgical techniques of graft fixation in anterior cruciate ligament reconstruction. *The American Journal of Sports Medicine*, 15(3), 225-229.
- Kvist, J., Ek, A., Sporrstedt, K., & Good, L. (2005). Fear of re-injury: a hindrance for returning to sports after anterior cruciate ligament reconstruction. *Knee Surgery, Sports Traumatology, Arthroscopy*, 13(5), 393-397.
- Lam, K. C. (2016). *Integrating Patient-Rated Outcome Measures into Routine Patient Care: Overcoming Common Challenges and Barriers*. Paper presented at the National Athletic Trainers Association 67th Clinical Symposia & AT Expo, Baltimore, MD. <http://convention.nata.org/st-abstracts/>
- Lam, K. C., Thomas, S. S., Valier, A. R. S., McLeod, T. C. V., & Bay, C. (2017). Previous knee injury and health-related quality of life in collegiate athletes. *Journal of Athletic Training*, 52(6), 534-540.
- Lam, K. C., Valier, A. R. S., Bay, C. R., & McLeod, T. C. V. (2013). A unique patient population? Health-related quality of life in adolescent athletes versus general, healthy adolescent individuals. *Journal of Athletic Training*, 48(2), 233-241. doi:10.4085/1062-6050-48.2.12
- Larson, R. (1974). Rating sheet for knee function. In I. S. Smillie (Ed.), *Disease of the knee joint* (pp. 29-30). Edinburgh: Churchill Livingstone.

- Lee, D. Y., Karim, S. A., & Chang, H. C. (2008). Return to sports after anterior cruciate ligament reconstruction--a review of patients with minimum 5-year follow-up. *Annals Academy of Medicine Singapore*, 37(4), 273.
- Leong, K. P., Yeak, S. C., Saurajen, A. S., Mok, P. K., Earnest, A., Siow, J. K., . . . Chng, H. H. (2005). Why generic and disease-specific quality-of-life instruments should be used together for the evaluation of patients with persistent allergic rhinitis. *Clinical & Experimental Allergy*, 35(3), 288-298. doi:10.1111/j.1365-2222.2005.02201.x
- Linacre, J. M. (2002). Optimizing rating scale category effectiveness. *Journal of Applied Measurement*, 3(1), 85-106.
- Lohmander, L., Östenberg, A., Englund, M., & Roos, H. (2004). High prevalence of knee osteoarthritis, pain, and functional limitations in female soccer players twelve years after anterior cruciate ligament injury. *Arthritis & Rheumatism*, 50(10), 3145-3152.
- Lohmander, L. S., Englund, P. M., Dahl, L. L., & Roos, E. M. (2007). The long-term consequence of anterior cruciate ligament and meniscus injuries osteoarthritis. *The American Journal of Sports Medicine*, 35(10), 1756-1769.
- Lohr, K. N. (1992). Applications of health status assessment measures in clinical practice: overview of the Third Conference on Advances in Health Status Assessment. *Medical Care*, 30(5), MS1-MS14.
- Lord, F. M. (1953). The relation of test score to the trait underlying the test. *Educational Testing Service Research Report Series*, 13, 517-549.

Lord, F. M. (1977). Practical applications of item characteristic curve theory*.

Journal of Educational Measurement, 14(2), 117-138.

Lord, F. M. (1980). *Applications of item response theory to practical testing problems*. New York: Routledge.

Lord, F. M., Novick, M. R., & Birnbaum, A (1968). *Statistical theories of mental test scores*. Charlotte: Information Age Publishing

Lukianov, A. V., Gillquist, J., Grana, W. A., & Dehaven, K. E. (1987). An anterior cruciate ligament (ACL) evaluation format for assessment of artificial or autologous anterior cruciate reconstruction results. *Clinical Orthopaedics and Related Research, 218*, 167-180.

Lysholm, J., & Gillquist, J. (1982). Evaluation of knee ligament surgery results with special emphasis on use of a scoring scale. *The American Journal of Sports Medicine, 10*(3), 150-154.

Lysholm, J., & Tegner, Y. (2007). Knee injury rating scales. *Acta Orthopaedica, 78*(4), 445-453.

Maffulli, N., Longo, U. G., Gougoulas, N., Loppini, M., & Denaro, V. (2010). Long-term health outcomes of youth sports injuries. *British Journal of Sports Medicine, 44*(1), 21-25.

Maly, M. R., Costigan, P. A., & Olney, S. J. (2006). Determinants of self-report outcome measures in people with knee osteoarthritis. *Archives of Physical Medicine and Rehabilitation, 87*(1), 96-104.

- Marais, I., & Andrich, D. (2008). Formalizing dimension and response violations of local independence in the unidimensional Rasch model. *Journal of Applied Measurement, 9*(3), 200-215.
- Marra, C. A., Woolcott, J. C., Kopec, J. A., Shojania, K., Offer, R., Brazier, J. E., . . . Anis, A. H. (2005). A comparison of generic, indirect utility measures (the HUI2, HUI3, SF-6D, and the EQ-5D) and disease-specific instruments (the RAQoL and the HAQ) in rheumatoid arthritis. *Social Science & Medicine, 60*(7), 1571-1582.
- Marshall, J. L., Fetto, J. F., & Botero, P. M. (1977). Knee ligament injuries: a standardized evaluation method. *Clinical Orthopaedics and Related Research, 123*, 115-129.
- Marshall, S. W., Covassin, T., Dick, R., Nassar, L. G., & Agel, J. (2007). Descriptive epidemiology of collegiate women's gymnastics injuries: National Collegiate Athletic Association Injury Surveillance System, 1988-1989 through 2003-2004. *Journal of Athletic Training, 42*(2), 234-240.
- Marshall, S. W., Hamstra-Wright, K. L., Dick, R., Grove, K. A., & Agel, J. (2007). Descriptive epidemiology of collegiate women's softball injuries: National Collegiate Athletic Association Injury Surveillance System, 1988–1989 through 2003–2004. *Journal of Athletic Training, 42*(2), 286-294.
- Marx, R. G., Jones, E. C., Allen, A. A., Altchek, D. W., O'Brien, S. J., Rodeo, S. A., . . . Wickiewicz, T. L. (2001). Reliability, validity, and responsiveness of four knee outcome scales for athletic patients. *The Journal of Bone and Joint Surgery American Volume, 83*(10), 1459-1469.

- Marx, R. G., Stump, T. J., Jones, E. C., Wickiewicz, T. L., & Warren, R. F. (2001). Development and evaluation of an activity rating scale for disorders of the knee. *The American Journal of Sports Medicine*, 29(2), 213-218.
- Masala, C., & Petretto, D. R. (2008). From disablement to enablement: conceptual models of disability in the 20th century. *Disability and Rehabilitation*, 30(17), 1233-1244.
- Masters, E. R. (1974). The relationship between number of response categories and reliability of Likert-type questionnaires. *Journal of Educational Measurement*, 11(1), 49-53. doi:10.1111/j.1745-3984.1974.tb00970.x
- Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika*, 47(2), 149-174.
- Matell, M. S., & Jacoby, J. (1971). Is there an optimal number of alternatives for Likert scale items? Study I: Reliability and validity. *Educational and Psychological Measurement*, 31(3), 657-674.
- McHorney, C. A. (1999). Health status assessment methods for adults: Past accomplishments and future challenges 1. *Annual Review of Public Health*, 20(1), 309-335.
- McHorney, C. A., Ware Jr, J. E., Rogers, W., Raczek, A. E., & Lu, J. R. (1992). The validity and relative precision of MOS short-and long-form health status scales and Dartmouth COOP charts: results from the Medical Outcomes Study. *Medical Care*, MS253-MS265.

- McLeod, T. C. V., Snyder, A. R., Parsons, J. T., Curtis Bay, R., Michener, L. A., & Sauers, E. L. (2008). Using disablement models and clinical outcomes assessment to enable evidence-based athletic training practice, part II: clinical outcomes assessment. *Journal of Athletic Training, 43*(4), 437-445.
- Miner, A. L., Lingard, E. A., Wright, E. A., Sledge, C. B., & Katz, J. N. (2003). Knee range of motion after total knee arthroplasty: how important is this as an outcome measure? *The Journal of Arthroplasty, 18*(3), 286-294.
- Muaidi, Q., Nicholson, L., & Refshauge, K. (2009). Do elite athletes exhibit enhanced proprioceptive acuity, range and strength of knee rotation compared with non-athletes? *Scandinavian Journal of Medicine & Science in Sports, 19*(1), 103-112.
- Nagi, S. (1965). Some conceptual issues in disability and rehabilitation. In M. Sussman (Ed.), *Sociology and rehabilitation* (pp. 100-113). Washington, DC: American Sociological Association.
- Nagi, S. Z. (1991). Disability concepts revisited: implications for prevention. *Disability in America: Toward a national agenda for prevention, 309-327*.
- Nandakumar, R. (1994). Assessing dimensionality of a set of item responses-comparison of different approaches. *Journal of Educational Measurement, 31*(1), 17-35.
- NATA Executive Committee for Education. (2011). *Athletic Training Education Competencies*. Retrieved from <http://caate.net/wp-content/uploads/2014/06/5th-Edition-Competencies.pdf>

- National Collegiate Athletic Association. (February 24th, 2012). Changes approved for 2012 season. Retrieved from <http://www.ncaa.com/news/football/article/2012-02-24/changes-approved-2012-season>
- National Collegiate Athletic Association Sports Science Institute. (October 16, 2014). Sports Injuries. Retrieved from <http://www.ncaa.org/health-and-safety/medical-conditions/sports-injuries>
- Noyes, F., Matthews, D., Mooar, P., & Grood, E. (1983). The symptomatic anterior cruciate-deficient knee. Part II: the results of rehabilitation, activity modification, and counseling on functional disability. *The Journal of Bone and Joint Surgery American Volume*, 65(2), 163-174.
- Noyes, F. R., & Barber-Westin, S. D. (1997). A comparison of results in acute and chronic anterior cruciate ligament ruptures of arthroscopically assisted autogenous patellar tendon reconstruction. *The American Journal of Sports Medicine*, 25(4), 460-471.
- Noyes, F. R., Barber, S. D., & Mooar, L. A. (1989). A rationale for assessing sports activity levels and limitations in knee disorders. *Clinical Orthopaedics and Related Research*, 246, 238-249.
- Noyes, F. R., Mooar, P., Matthews, D., & Butler, D. (1983). The symptomatic anterior cruciate-deficient knee. Part I: the long-term functional disability in athletically active individuals. *The Journal of Bone and Joint Surgery American Volume*, 65(2), 154-162.

Odensten, M., Hamberg, P., Nordin, M., Lysholm, J., & Gillquist, J. (1985).

Surgical or Conservative Treatment of the Acutely Torn Anterior Cruciate Ligament: A Randomized Study with Short-term Follow-up Observations. *Clinical Orthopaedics and Related Research*, 198, 87-93.

Oretorp, N., Gillquist, J., & Liljedahl, S.-O. (1979). Long term results of surgery for non-acute anteromedial rotatory instability of the knee. *Acta Orthopaedica Scandinavica*, 50(3), 329-336.

Paterno, M. V., Ford, K. R., Myer, G. D., Heyl, R., & Hewett, T. E. (2007). Limb asymmetries in landing and jumping 2 years following anterior cruciate ligament reconstruction. *Clinical Journal of Sport Medicine*, 17(4), 258-262. doi:10.1097/JSM.0b013e31804c77ea

The Patient Protection and Affordable Care Act, 111-148 C.F.R. (2010).

Paxton, E. W., Fithian, D. C., Stone, M. L., & Silva, P. (2003). The reliability and validity of knee-specific and general health instruments in assessing acute patellar dislocation outcomes. *The American Journal of Sports Medicine*, 31(4), 487-492.

Physical Activity Council. (2016). *The Physical Activity Council's annual study tracking sports, fitness, and recreation participation in the US*. Retrieved from <http://www.physicalactivitycouncil.com/PDFs/current.pdf>

Physical Activity Guidelines Advisory Committee. (2008). Physical activity guidelines advisory committee report, 2008. *Washington, DC: US Department of Health and Human Services, 2008, A1-H14*.

- Ragan, B. G., Herrmann, S. D., Kang, M., & Mack, M. G. (2009). Psychometric evaluation of the Standardized Assessment of Concussion: evaluation of baseline score validity using item analysis. *Athletic Training and Sports Health Care, 1*(4), 180-187.
- Rasch, G. (1961). *On general laws and the meaning of measurement in psychology*. Paper presented at the Proceedings of the fourth Berkeley symposium on mathematical statistics and probability.
- Reckase, M. D. (1979). Unifactor latent trait models applied to multifactor tests: Results and implications. *Journal of Educational and Behavioral Statistics, 4*(3), 207-230.
- Reeve, B. B., Hays, R. D., Bjorner, J. B., Cook, K. F., Crane, P. K., Teresi, J. A., . . . Hambleton, R. K. (2007). Psychometric evaluation and calibration of health-related quality of life item banks: plans for the Patient-Reported Outcomes Measurement Information System (PROMIS). *Medical Care, 45*(5), S22-S31.
- Roos, E. M., Roos, H. P., Lohmander, L. S., Ekdahl, C., & Beynnon, B. D. (1998). Knee Injury and Osteoarthritis Outcome Score (KOOS)—development of a self-administered outcome measure. *Journal of Orthopaedic & Sports Physical Therapy, 28*(2), 88-96.
- Roos, E. M., & Toksvig-Larsen, S. (2003). Knee injury and Osteoarthritis Outcome Score (KOOS)-validation and comparison to the WOMAC in total knee replacement. *Health and Quality of Life Outcomes, 1*(1), 17.

- Rosenbaum, P., & Stewart, D. (2004). *The World Health Organization International Classification of Functioning, Disability, and Health: a model to guide clinical thinking, practice and research in the field of cerebral palsy*. Paper presented at the Seminars in pediatric neurology.
- Rosenbaum, P. R. (1984). Testing the conditional independence and monotonicity assumptions of item response theory. *Psychometrika*, 49(3), 425-435.
- Roter, D. L. (1977). Patient participation in the patient-provider interaction: the effects of patient question asking on the quality of interaction, satisfaction and compliance. *Health Education & Behavior*, 5(4), 281-315.
- Rothwell, P., McDowell, Z., Wong, C., & Dorman, P. (1997). Doctors and patients don't agree: cross sectional study of patients' and doctors' perceptions and assessments of disability in multiple sclerosis. *British Journal of Sports Medicine*, 314(7094), 1580-1583.
- Safrit, M. J. (1987). The applicability of item response theory to tests of motor behavior. *Research Quarterly for Exercise and Sport*, 58(3), 213-215.
- Schneidert, M., Hurst, R., Miller, J., & Üstün, B. (2003). The role of environment in the International Classification of Functioning, Disability and Health (ICF). *Disability and Rehabilitation*, 25(11-12), 588-595.
- Selby, J. V., Beal, A. C., & Frank, L. (2012). The Patient-Centered Outcomes Research Institute (PCORI) national priorities for research and initial research agenda. *Journal of the American Medical Association*, 307(15), 1583-1584.

- Sernert, N., Kartus, J., Köhler, K., Stener, S., Larsson, J., Eriksson, B. I., & Karlsson, J. (1999). Analysis of subjective, objective and functional examination tests after anterior cruciate ligament reconstruction A follow-up of 527 patients. *Knee Surgery, Sports Traumatology, Arthroscopy*, 7(3), 160-165.
- Sgaglione, N. A., Del Pizzo, W., Fox, J. M., & Friedman, M. J. (1995). Critical analysis of knee ligament rating systems. *The American Journal of Sports Medicine*, 23(6), 660-667.
- Shirazi, C. P., Israel, H. A., & Kaar, S. G. (2016). Is the Marx Activity Rating Scale reliable in patients younger than 18 years? *Sports Health*, 8(2), 145-147.
- Sireci, S. G., Thissen, D., & Wainer, H. (1991). On the reliability of testlet-based tests. *Journal of Educational Measurement*, 28(3), 237-247.
- Snyder, A. R., Martinez, J. C., Bay, R. C., Parsons, J. T., Sauers, E. L., & McLeod, T. C. V. (2010). Health-related quality of life differs between adolescent athletes and adolescent nonathletes. *Journal of Sport Rehabilitation*, 19(3), 237-248.
- Snyder, A. R., Parsons, J. T., Valovich McLeod, T. C., Curtis Bay, R., Michener, L. A., & Sauers, E. L. (2008). Using disablement models and clinical outcomes assessment to enable evidence-based athletic training practice, part I: disablement models. *Journal of Athletic Training*, 43(4), 428-436.
- Solonen, K. A., & Rokkanen, P. (1967). Operative treatment of torn ligaments in injuries of the knee joint. *Acta Orthopaedica Scandinavica*, 38(1-4), 67-80.

- Spray, J. A. (1987). Recent developments in measurement and possible applications to the measurement of psychomotor behavior. *Research Quarterly for Exercise and Sport*, 58(3), 203-209.
- Stout, W. (1987). A nonparametric approach for assessing latent trait unidimensionality. *Psychometrika*, 52(4), 589-617.
- Straub, T., & Hunter, R. E. (1988). Acute anterior cruciate ligament repair. *Clinical Orthopaedics and Related Research*, 227, 238-250.
- Swiontkowski, M. F. (2005). Musculoskeletal outcomes measures and instruments. *The Journal of Bone and Joint Surgery American Volume*, 87(9), 2154-2154.
- Tagesson, S., Öberg, B., Good, L., & Kvist, J. (2008). A comprehensive rehabilitation program with quadriceps strengthening in closed versus open kinetic chain exercise in patients with anterior cruciate ligament deficiency: a randomized clinical trial evaluating dynamic tibial translation and muscle function. *The American Journal of Sports Medicine*, 36(2), 298-307.
- Tegner, Y., & Lysholm, J. (1985). Rating systems in the evaluation of knee ligament injuries. *Clinical Orthopaedics and Related Research*, 198, 42-49.
- Terwee, C. B., Mokkink, L. B., Steultjens, M. P. M., & Dekker, J. (2006). Performance-based methods for measuring the physical function of patients with osteoarthritis of the hip or knee: a systematic review of

measurement properties. *Rheumatology*, 45(7), 890-902.

doi:10.1093/rheumatology/kei267

Thissen, D., Steinberg, L., & Mooney, J. A. (1989). Trace lines for testlets: A use of multiple-categorical-response models. *Journal of Educational Measurement*, 26(3), 247-260.

Union of the Physically Impaired Against Segregation. (1975). *Fundamental principles of disability*. London: UPIAS.

Urwin, M., Symmons, D., Allison, T., Brammah, T., Busby, H., Roxby, M., . . .

Williams, G. (1998). Estimating the burden of musculoskeletal disorders in the community: the comparative prevalence of symptoms at different anatomical sites, and the relation to social deprivation. *Annals of the Rheumatic Diseases*, 57(11), 649-655.

Valier, A. R. S., Jennings, A. L. J., Parsons, J. T., & Vela, L. I. (2014). Benefits of and barriers to using patient-rated outcome measures in athletic training. *Journal of Athletic Training*, 49(5), 674-683. doi:10.4085/1062-6050-49.3.15

van der List, J. P., & DiFelice, G. S. (2017). Range of motion and complications following primary repair versus reconstruction of the anterior cruciate ligament. *Knee*, 24(4), 798-807. doi:10.1016/j.knee.2017.04.007

Verrall, G. M., Kalairajah, Y., Slavotinek, J., & Spriggins, A. (2006). Assessment of player performance following return to sport after hamstring muscle strain injury. *Journal of Science and Medicine in Sport*, 9(1), 87-90.

- Wainer, H., & Thissen, D. (1996). How is reliability related to the quality of test scores? What is the effect of local dependence on reliability? *Educational Measurement: Issues and Practice*, 15(1), 22-29.
- Wang, D., Jones, M. H., Khair, M. M., & Miniaci, A. (2010). Patient-reported outcome measures for the knee. *The Journal of Knee Surgery*, 23(3), 137-152.
- Webb, R., Brammah, T., Lunt, M., Urwin, M., Allison, T., & Symmons, D. (2004). Opportunities for prevention of 'clinically significant' knee pain: results from a population-based cross sectional survey. *Journal of Public Health*, 26(3), 277-284.
- Wiger, P., Brandsson, S., Kartus, J., Eriksson, B., & Karlsson, J. (1999). A comparison of results after arthroscopic anterior cruciate ligament reconstruction in female and male competitive athletes A two-to five-year follow-up of 429 patients. *Scandinavian Journal of Medicine & Science in Sports*, 9(5), 290-295.
- Wilson, M. (1992). The ordered partition model: An extension of the partial credit model. *Applied Psychological Measurement*, 16, 309-325.
- Wood, T. M. (1987). Putting item response theory into perspective. *Research Quarterly for Exercise and Sport*, 58(3), 216-220.
- World Health Organization. (1980). International classification of impairments, disabilities, and handicaps: a manual of classification relating to the consequences of disease, published in accordance with resolution WHA29. 35 of the Twenty-ninth World Health Assembly, May 1976.

- World Health Organization. (2001). *International Classification of Functioning, Disability and Health*. Geneva: World Health Organization.
- World Health Organization. (2016). *International Classification of Functioning, Disability and Health*. Geneva: World health Organization.
- Wright, B. D., & Douglas, G. A. (1977). Best procedures for sample-free item analysis. *Applied Psychological Measurement*, 1(2), 281-295.
- Wright, B. D., Linacre, J. M., Gustafson, J., & Martin-Lof, P. (1994). Reasonable mean-square fit values. *Rasch Measurement Transactions*, 8(3), 370.
- Yamazaki, J., Muneta, T., Ju, Y. J., & Sekiya, I. (2009). Differences in kinematics of single leg squatting between anterior cruciate ligament-injured patients and healthy controls. *Knee Surgery, Sports Traumatology, Arthroscopy*, 18(1), 56. doi:10.1007/s00167-009-0892-z
- Yeh, P. C., Starkey, C., Lombardo, S., Vitti, G., & Kharrazi, F. D. (2012). Epidemiology of isolated meniscal injury and its effect on performance in athletes from the National Basketball Association. *The American Journal of Sports Medicine*, 40(3), 589-594.
- Young, K., White, P., & McTeer, W. (1994). Body talk: Male athletes reflect on sport, injury, and pain. *Sociology of Sport Journal*, 11(2), 175-194.

APPENDICES

APPENDIX A: IRB APPROVAL LETTER CHAPTER III

IRB

INSTITUTIONAL REVIEW BOARD

Office of Research Compliance,
010A Sam Ingram Building,
2269 Middle Tennessee Blvd
Murfreesboro, TN 37129



IRBN007 – EXEMPTION DETERMINATION NOTICE

Friday, June 17, 2016

Investigator(s): James L Farnsworth II (PI), and Dr. Minsoo Kang (FA)
Investigator(s) Email(s): jlf6g@mtmail.mtsu.edu
Department: Health and Human Performance

Study Title: Psychometric Evaluation of Patient Reported Knee Outcomes for
Assessment of Knee-Related Function
Protocol ID: 16-1291

Dear Investigator(s),

The above identified research proposal has been reviewed by the MTSU Institutional Review Board (IRB) through the **EXEMPT** review mechanism under 45 CFR 46.101(b)(2) within the research category (2) *Educational Tests*. A summary of the IRB action and other particulars in regard to this protocol application is tabulated as shown below:

IRB Action	EXEMPT from further IRB review***	
Date of expiration	NOT APPLICABLE	
Sample Size	160 (ONE HUNDRED AND SIXTY)	
Participant Pool	Adult participants (18 yrs or older) who have a history of knee injury	
Mandatory Requirements	No identifiable data may be collected and only documents approved by the IRB can be used in this study.	
Additional Restrictions	Must collect informed consent	
Comments	N/A	
Amendments	Date N/A	Post-Approval Amendments N/A

***This exemption determination only allows above defined protocol from further IRB review such as continuing review. However, the following post-approval requirements still apply:

- Addition/removal of subject population should not be implemented without IRB approval
- Change in investigators must be notified and approved
- Modifications to procedures must be clearly articulated in an addendum request and the proposed changes must not be incorporated without an approval
- Be advised that the proposed change must comply within the requirements for exemption
- Changes to the research location must be approved – appropriate permission letter(s) from external institutions must accompany the addendum request form
- Changes to funding source must be notified via email (irb_submissions@mtsu.edu)

- The exemption does not expire as long as the protocol is in good standing
- Project completion must be reported via email (irb_submissions@mtsu.edu)
- Research-related injuries to the participants and other events must be reported within 48 hours of such events to compliance@mtsu.edu

The current MTSU IRB policies allow the investigators to make the following types of changes to this protocol without the need to report to the Office of Compliance, as long as the proposed changes do not result in the cancellation of the protocols eligibility for exemption:

- Editorial and minor administrative revisions to the consent form or other study documents
- Increasing/decreasing the participant size

The investigator(s) indicated in this notification should read and abide by all applicable post-approval conditions imposed with this approval. [Refer to the post-approval guidelines posted in the MTSU IRB's website](#). Any unanticipated harms to participants or adverse events must be reported to the Office of Compliance at (615) 494-8918 within 48 hours of the incident.

All of the research-related records, which include signed consent forms, current & past investigator information, training certificates, survey instruments and other documents related to the study, must be retained by the PI or the faculty advisor (if the PI is a student) at the secure location mentioned in the protocol application. The data storage must be maintained for at least three (3) years after study completion. Subsequently, the researcher may destroy the data in a manner that maintains confidentiality and anonymity. IRB reserves the right to modify, change or cancel the terms of this letter without prior notice. Be advised that IRB also reserves the right to inspect or audit your records if needed.

Sincerely,

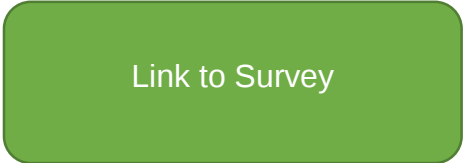
Institutional Review Board
Middle Tennessee State University

Quick Links:

[Click here](#) for a detailed list of the post-approval responsibilities.
More information on exmpt procedures can be found [here](#).

APPENDIX B: RECRUITMENT SCRIPT CHAPTER III

My name is James Farnsworth. I am a doctoral candidate at MTSU and I am conducting research on patient reported outcomes for evaluating knee function. Would you be willing to take approximately 20 minutes to complete a brief questionnaire? Your responses will be anonymous. Your participation is completely voluntary. If you have any questions about the questionnaire, please feel free to contact me via email at jlf6g@mtmail.mtsu.edu. If you choose to participate, please click the link below to begin the questionnaire. Thank you in advance for your time.



Link to Survey

APPENDIX C: COMBINED KNEE FUNCTION ASSESSMENT TOOL (KNEE INJURY AND OSTEOARTHRITIS OUTCOME SCORE, INTERNATIONAL KNEE DOCUMENTATION COMMITTEE SUBJECTIVE KNEE FORM, MARX ACTIVITY RATING SCALE)

6/15/2016

Comprehensive Knee Function Evaluation Tool

Comprehensive Knee Function Evaluation Tool

Description of Study

Purpose:

The purpose of this study is to investigate the quality of three commonly used surveys for measuring knee-related function. The three instruments include the Knee Injury Osteoarthritis Outcome Score (KOOS), the International Knee Documentation Committee Subjective Knee Form (IKDC), and the Marx Activity Scale.

Procedures:

You are being asked to participate in this study and complete each of these three surveys as well as a brief injury history form. Each of these three instruments include items related to knee pain, symptoms, ability to complete daily living tasks (such as walking up and down stairs), ability to perform sports-related activities, and overall quality of life. Combined the three instruments include 65 questions which should take approximately 10 to 15 minutes to complete. The knee is one of the most commonly injured regions of the body in sports. Following injury it is important to be able to measure a patient's knee function to determine when they are able to return to play. Your participation in this study will be used to help us to evaluate and improve these instruments, thus providing better quality of care to those who have been injured.

Risks/Benefits:

The goal of this study is to obtain a more complete understanding of how well these instruments perform in a variety of adult patients (both athletic and non-athletic). The anticipated risk for this study are minimal (e.g., have you experienced any difficulties performing the following activities: ascending stairs, getting on/off the toilet, etc.). To minimize the potential risk to your privacy, your questionnaire data will remain anonymous.

Confidentiality:

To minimize the risk to your privacy, we will not be collecting any data containing personal identifiers. All data collected for this study will be stored electronically on a password protected server. Only the study investigators will have access to the data.

Principle Investigator/ Contact Information:

James L. Farnsworth II
Email: jlf6q@mtmail.mtsu.edu

Informed Consent:

6/15/2016

Comprehensive Knee Function Evaluation Tool

Participating in this project is voluntary, and refusal to participate or withdrawing from participation at any time during the project will involve no penalty or loss of benefits to which you might otherwise be entitled. All efforts, within reason, will be made to keep the personal information in your research record private but total privacy cannot be promised, for example, your information may be shared with the Middle Tennessee State University Institutional Review Board. In the event of questions or difficulties of any kind during or following participation, you may contact the Principle Investigator as indicated above. For additional information about giving consent or your rights as a participant in this study, please feel free to contact the MTSU Office of Compliance at (615) 494-8919

I have read the above information and my questions have been answered satisfactorily by the project staff. I believe I understand the purpose, benefits, and risks of the study. By clicking the link at the bottom of the page below I give my informed and free consent to be a participant.

Demographic Information

This section will be used to collect basic information related to your personal characteristics (age, height, weight, etc.) and your history (or lack of) knee-related injuries.

1. Age

2. Sex

Mark only one oval.

☐ Male

☐ Female

3. Height

4. Weight

5. Occupation

Physical Activity Assessment

The purpose of this section is to determine your level of physical activity. Use the following table below as a reference when determining your level of usual activity.

Examples of Moderate and Vigorous Activities

Examples of Moderate Intensity Physical Activity	Examples of Vigorous Intensity Physical Activity
- Walking briskly (3 miles per hour or faster) - Bicycling slower than 10 miles per hour - Water aerobics - Tennis (doubles)	- Race-walking, jogging, or running - Swimming laps - Jumping rope - Hiking uphill or with a heavy backpack

6/15/2016

Comprehensive Knee Function Evaluation Tool

6. Read all four options carefully; then select the option that best describes your usual weekly activity.

Mark only one oval.

- ☐ Inactive - May perform some light physical activity (such as walking, climbing stairs), but no moderate or vigorous physical activity)
- ☐ Lightly Active - Some moderate or vigorous intensity physical activity, but less than 150 minutes per week of moderate intensity physical activity, or 75 minutes of vigorous intensity physical activity
- ☐ Moderately Active - 150 to 300 minutes per week of moderate intensity physical activity, 75 to 150 minutes of vigorous intensity physical activity, or a combination of the two
- ☐ Highly Active - More than 300 minutes per week of moderate intensity physical activity, 150 minutes of vigorous intensity physical activity, or a combination of the two

7. Do you currently participate in recreational/organized sports?

Mark only one oval.

- ☐ Yes
- ☐ No

8. If you answered yes to the previous question (sports participation) please provide a brief description of your level of engagement in these activities

Injury History

This section will ask you questions about any current knee injuries. If you do not have any knee injuries. Answer "no" to the first question; then skip to the next section by clicking the link at the bottom of the page.

9. Are you currently experiencing any knee-related injuries or disabilities.

Mark only one oval.

- ☐ Yes
- ☐ No

10. Provide a brief description of your injury

6/15/2016

Comprehensive Knee Function Evaluation Tool

11. What side of your body was/is the injury located on*Mark only one oval.*

- ☐ Left
- ☐ Right
- ☐ Both

12. Are you receiving treatment for your pain (therapy, surgery, etc).*Mark only one oval.*

- ☐ Yes
- ☐ No

13. Days since injury occurred:

14. Days since surgery occurred:

15. Do you currently have any other injuries/conditions that may limit your function not related to the knee?*Mark only one oval.*

- ☐ Yes
- ☐ No

Assessment Tool 1**Knee Injury and Osteoarthritis Outcome Score**

Instructions:

Please read each of the following questions below and circle the answer that best describes how you feel about your knee. If you are unsure about how to answer a question, please give the best answer that you can. Please answer all questions.

Symptoms:

These questions should be answered thinking of your knee symptoms during the last week

6/15/2016

Comprehensive Knee Function Evaluation Tool

16. Do you have swelling in your knee?*Mark only one oval.*

- ☐ Never
☐ Rarely
☐ Sometimes
☐ Often
☐ Always

17. Do you feel grinding, hear clicking or any other type of noise when your knee moves?*Mark only one oval.*

- ☐ Never
☐ Rarely
☐ Sometimes
☐ Often
☐ Always

18. Does your knee catch or hang up when moving?*Mark only one oval.*

- ☐ Never
☐ Rarely
☐ Sometimes
☐ Often
☐ Always

19. Can you straighten your knee fully?*Mark only one oval.*

- ☐ Never
☐ Rarely
☐ Sometimes
☐ Often
☐ Always

20. Can you bend your knee fully?*Mark only one oval.*

- ☐ Never
☐ Rarely
☐ Sometimes
☐ Often
☐ Always

Assessment Tool 1

6/15/2016

Comprehensive Knee Function Evaluation Tool

Knee Injury and Osteoarthritis Outcome Score

Instructions:

Please read each of the following questions below and circle the answer that best describes how you feel about your knee. If you are unsure about how to answer a question, please give the best answer that you can. Please answer all questions.

Stiffness:

The following questions concern the amount of joint stiffness you have experienced during the last week in your knee. Stiffness is a sensation of restriction or slowness in the ease with which you move your knee joint.

21. How severe is your knee joint stiffness after first waking in the morning?

Mark only one oval.

- ☐ None
☐ Mild
☐ Moderate
☐ Severe
☐ Extreme

22. How severe is your knee stiffness after sitting, lying, or resting later in the day?

Mark only one oval.

- ☐ None
☐ Mild
☐ Moderate
☐ Severe
☐ Extreme

Assessment Tool 1

Knee Injury and Osteoarthritis Outcome Score

Instructions:

Please read each of the following questions below and circle the answer that best describes how you feel about your knee. If you are unsure about how to answer a question, please give the best answer that you can. Please answer all questions.

Pain:

6/15/2016

Comprehensive Knee Function Evaluation Tool

23. How often do you experience knee pain?*Mark only one oval.*

- ☐ Never
☐ Monthly
☐ Weekly
☐ Daily
☐ Always

Pain related to specific activities:

What amount of knee pain have you experienced the last week during the following activities?

24. Twisting/pivoting on your knee*Mark only one oval.*

- ☐ None
☐ Mild
☐ Moderate
☐ Severe
☐ Extreme

25. Straightening knee fully*Mark only one oval.*

- ☐ None
☐ Mild
☐ Moderate
☐ Severe
☐ Extreme

26. Bending knee fully*Mark only one oval.*

- ☐ None
☐ Mild
☐ Moderate
☐ Severe
☐ Extreme

6/15/2016

Comprehensive Knee Function Evaluation Tool

27. Walking on flat surface*Mark only one oval.*

- ☐ None
☐ Mild
☐ Moderate
☐ Severe
☐ Extreme

28. Going up or down stairs*Mark only one oval.*

- ☐ None
☐ Mild
☐ Moderate
☐ Severe
☐ Extreme

29. At night while in bed*Mark only one oval.*

- ☐ None
☐ Mild
☐ Moderate
☐ Severe
☐ Extreme

30. Sitting or lying*Mark only one oval.*

- ☐ None
☐ Mild
☐ Moderate
☐ Severe
☐ Extreme

31. Standing upright*Mark only one oval.*

- ☐ None
☐ Mild
☐ Moderate
☐ Severe
☐ Extreme

Assessment Tool 1

6/15/2016

Comprehensive Knee Function Evaluation Tool

Knee Injury and Osteoarthritis Outcome Score

Instructions:

Please read each of the following questions below and circle the answer that best describes how you feel about your knee. If you are unsure about how to answer a question, please give the best answer that you can. Please answer all questions.

Function, daily living

The following questions concern your physical function. By this we mean your ability to move around and to look after yourself. For each of the following activities please indicate the degree of difficulty you have experienced in the last week due to your knee.

32. Descending stairs

Mark only one oval.

- ☐ None
☐ Mild
☐ Moderate
☐ Severe
☐ Extreme

33. Ascending stairs

Mark only one oval.

- ☐ None
☐ Mild
☐ Moderate
☐ Severe
☐ Extreme

34. Rising from sitting

Mark only one oval.

- ☐ None
☐ Mild
☐ Moderate
☐ Severe
☐ Extreme

6/15/2016

Comprehensive Knee Function Evaluation Tool

35. Standing*Mark only one oval.*

- ☐ None
☐ Mild
☐ Moderate
☐ Severe
☐ Extreme

36. Bending to floor/pick up an object*Mark only one oval.*

- ☐ None
☐ Mild
☐ Moderate
☐ Severe
☐ Extreme

37. Walking on flat surface*Mark only one oval.*

- ☐ None
☐ Mild
☐ Moderate
☐ Severe
☐ Extreme

38. Going shopping*Mark only one oval.*

- ☐ None
☐ Mild
☐ Moderate
☐ Severe
☐ Extreme

39. Putting on socks/stockings*Mark only one oval.*

- ☐ None
☐ Mild
☐ Moderate
☐ Severe
☐ Extreme

6/15/2016

Comprehensive Knee Function Evaluation Tool

40. Rising from bed*Mark only one oval.*

- ☐ None
☐ Mild
☐ Moderate
☐ Severe
☐ Extreme

41. Taking off socks/stockings*Mark only one oval.*

- ☐ None
☐ Mild
☐ Moderate
☐ Severe
☐ Extreme

42. Lying in bed (turning over, maintaining knee position)*Mark only one oval.*

- ☐ None
☐ Mild
☐ Moderate
☐ Severe
☐ Extreme

43. Getting in/out of bath*Mark only one oval.*

- ☐ None
☐ Mild
☐ Moderate
☐ Severe
☐ Extreme

44. Sitting*Mark only one oval.*

- ☐ None
☐ Mild
☐ Moderate
☐ Severe
☐ Extreme

6/15/2016

Comprehensive Knee Function Evaluation Tool

45. Getting on/off toilet*Mark only one oval.*

- ☐ None
- ☐ Mild
- ☐ Moderate
- ☐ Severe
- ☐ Extreme

46. Heavy domestic duties (moving heavy boxes, scrubbing floors, etc.)*Mark only one oval.*

- ☐ None
- ☐ Mild
- ☐ Moderate
- ☐ Severe
- ☐ Extreme

47. Light domestic duties (cooking, dusting, etc.)*Mark only one oval.*

- ☐ None
- ☐ Mild
- ☐ Moderate
- ☐ Severe
- ☐ Extreme

Assessment Tool 1**Knee Injury and Osteoarthritis Outcome Score**

Instructions:

Please read each of the following questions below and circle the answer that best describes how you feel about your knee. If you are unsure about how to answer a question, please give the best answer that you can. Please answer all questions.

Function, sports, and recreational activities

The following questions concern your physical function when being active on a higher level. The questions should be answered thinking of what degree of difficulty you have experienced during the last week due to your knee.

6/15/2016

Comprehensive Knee Function Evaluation Tool

48. Squatting*Mark only one oval.*

- ☐ None
☐ Mild
☐ Moderate
☐ Severe
☐ Extreme

49. Running*Mark only one oval.*

- ☐ None
☐ Mild
☐ Moderate
☐ Severe
☐ Extreme

50. Jumping*Mark only one oval.*

- ☐ None
☐ Mild
☐ Moderate
☐ Severe
☐ Extreme

51. Twisting/pivoting on your knee*Mark only one oval.*

- ☐ None
☐ Mild
☐ Moderate
☐ Severe
☐ Extreme

52. Kneeling*Mark only one oval.*

- ☐ None
☐ Mild
☐ Moderate
☐ Severe
☐ Extreme

Assessment Tool 1

6/15/2016

Comprehensive Knee Function Evaluation Tool

Knee Injury and Osteoarthritis Outcome Score

Instructions:

Please read each of the following questions below and circle the answer that best describes how you feel about your knee. If you are unsure about how to answer a question, please give the best answer that you can. Please answer all questions.

Quality of Life

53. How often are you aware of your knee problem?

Mark only one oval.

- ☐ Never
☐ Monthly
☐ Weekly
☐ Daily
☐ Always

54. Have you modified your life style to avoid potentially damaging activities to your knee?

Mark only one oval.

- ☐ Not at all
☐ Mildly
☐ Moderately
☐ Severely
☐ Totally

55. How much are you troubled with lack of confidence in your knee?

Mark only one oval.

- ☐ None
☐ Mild
☐ Moderate
☐ Severe
☐ Extreme

56. In general, how much difficulty do you have with your knee?

Mark only one oval.

- ☐ None
☐ Mild
☐ Moderate
☐ Severe
☐ Extreme

Assessment Tool 2

International Knee Documentation Committee Subjective Knee Evaluation Form

Symptoms

Grade symptoms at the highest activity level at which you think you could function without significant symptoms, even if you are not actually performing activities at this level. Mark your answer using the bubble provided for each item.

57. What is the highest level of activity that you can perform without significant knee pain?

Mark only one oval.

- ☐ Very strenuous activities like jumping or pivoting as in basketball or soccer
- ☐ Strenuous activities like heavy physical work; skiing or tennis
- ☐ Moderate activities like moderate physical work, running or jogging
- ☐ Light activities like walking, housework or yard work
- ☐ Unable to perform any of the above activities due to knee pain

58. During the past 4 weeks, or since your injury, how often have you had pain?

Mark only one oval.

	0	1	2	3	4	5	6	7	8	9	10	
Never	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	Constantly

59. If you have pain, how severe is it?

Mark only one oval.

	0	1	2	3	4	5	6	7	8	9	10	
No pain	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	Worse pain imaginable

60. During the past 4 weeks, or since your injury, how stiff or swollen was your knee?

Mark only one oval.

- ☐ Not at all
- ☐ Mildly
- ☐ Moderately
- ☐ Very
- ☐ Extremely

6/15/2016

Comprehensive Knee Function Evaluation Tool

61. **What is the highest level of activity that you can perform without significant swelling in your knee?**

Mark only one oval.

- ☐ Very strenuous activities like jumping or pivoting as in basketball or soccer
- ☐ Strenuous activities like heavy physical work; skiing or tennis
- ☐ Moderate activities like moderate physical work, running or jogging
- ☐ Light activities like walking, housework or yard work
- ☐ Unable to perform any of the above activities due to knee pain

62. **During the past 4 weeks, or since your injury, did your knee lock or catch?**

Mark only one oval.

- ☐ Yes
- ☐ No

63. **What is the highest level of activity that you can perform without significant giving way in your knee?**

Mark only one oval.

- ☐ Very strenuous activities like jumping or pivoting as in basketball or soccer
- ☐ Strenuous activities like heavy physical work; skiing or tennis
- ☐ Moderate activities like moderate physical work, running or jogging
- ☐ Light activities like walking, housework or yard work
- ☐ Unable to perform any of the above activities due to knee pain

Assessment Tool 2

International Knee Documentation Committee Subjective Knee Evaluation Form

Sports Activities:

64. **What is the highest level of activity you can participate in on a regular basis?**

Mark only one oval.

- ☐ Very strenuous activities like jumping or pivoting as in basketball or soccer
- ☐ Strenuous activities like heavy physical work; skiing or tennis
- ☐ Moderate activities like moderate physical work, running or jogging
- ☐ Light activities like walking, housework or yard work
- ☐ Unable to perform any of the above activities due to knee pain

6/15/2016

Comprehensive Knee Function Evaluation Tool

65. For each of the nine activities below, select the description that best describes your ability to perform each activity:

Mark only one oval per row.

	Not difficult at all	Minimally difficult	Moderately difficult	Extremely difficult	Unable to do
Go up stairs	☐	☐	☐	☐	☐
Go down stairs	☐	☐	☐	☐	☐
Kneel on the front of your knee	☐	☐	☐	☐	☐
Squat	☐	☐	☐	☐	☐
Sit with your knee bent	☐	☐	☐	☐	☐
Rise from a chair	☐	☐	☐	☐	☐
Run straight ahead	☐	☐	☐	☐	☐
Jump and land on your involved leg	☐	☐	☐	☐	☐
Stop and start quickly	☐	☐	☐	☐	☐

Assessment Tool 2

International Knee Documentation Committee Subjective Knee Evaluation Form

Function

How would you rate the function of your knee on a scale of 0 to 10 with 10 being normal, excellent function and 0 being the inability to perform any of your usual daily activities which may include sports?

66. Function prior to your knee injury

Mark only one oval.

	0	1	2	3	4	5	6	7	8	9	10	
Couldn't perform daily activities	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	No limitation in daily activities

67. Current function of your knee

Mark only one oval.

	0	1	2	3	4	5	6	7	8	9	10	
Couldn't perform daily activities	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	No limitation in daily activities

Assessment Tool 3

Marx Activity Scale

https://docs.google.com/forms/d/12vcPzltDw5BZl190kSY1pwCIIC7mebugd1dA_sEz13w/edit

17/18

6/15/2016

Comprehensive Knee Function Evaluation Tool

68. Please indicate how often you performed each activity in your healthiest and most active state, in the past year. Kindly select the option that best describes your level of activity

Mark only one oval per row.

	Less than one time in a month	One time in a month	One time in a week	2 or 3 times in a week	4 or more times in a week
Running: running while playing a sport or jogging	☐	☐	☐	☐	☐
Cutting: changing directions while running	☐	☐	☐	☐	☐
Deceleration: coming to a quick stop while running	☐	☐	☐	☐	☐
Pivoting: turning your body with your foot planted while playing sport; For example: skiing, skating, kicking, throwing, hitting a ball (glove, tennis, squash), etc.	☐	☐	☐	☐	☐

Powered by



APPENDIX D: IRB APPROVAL LETTER CHAPTER IV

March 6, 2017

Dear Mr. Farnsworth

IRB#: 2016-2017-021

Title of Proposal: **Development and Validation of an Improved Patient Reported Knee Outcomes Instrument**

This letter is to officially notify you of the approval of your project by the Buena Vista University IRB/IACUC Board. It is the Board's opinion that you have provided adequate safeguards for the rights and welfare of participants in this study. Your proposal is in compliance with the DHHS Regulations for the Protection of Human Subjects (45 CFR 46).

You are authorized to implement this study as of the Date of Final Approval: 03/06/2017.

This approval is Valid Until: 03/05/2018.

You must notify the IRB/IACUC Board immediately of any proposed changes to your research project. No changes are to be made without prior approval of the IRB/IACUC Board. You should report any unanticipated problems involving risks to the participants or others to the Board. For projects which continue beyond one year from the starting date,

APPENDIX E: CONSENT FORM CHAPTER IV

Informed Consent

Buena Vista University

Project Title: Development and validation of an improved patient-reported knee outcomes instrument

Purpose of Project: Patient-reported outcomes instruments are commonly used to assess function following injury. Although these instruments are commonly used many of the items on these instruments are not relevant to athletes making them less useful for monitoring recovery from injury. Recently we developed a new instrument to monitor recovery from knee-related injuries specifically for athletes, however, this instrument has not yet been tested. Therefore, the purpose of this study is to evaluate the quality of the new evaluation tool.

Procedures: You are being asked to participate in this study and complete a short questionnaire as well as a brief injury history form. The questionnaire includes items related to knee pain, symptoms, ability to complete daily living tasks (such as walking up and down stairs), ability to perform sports-related activities, and overall quality of life. The questionnaire should take approximately 15 minutes to complete. Additionally, you may be asked to provide feedback about the overall quality of the questionnaire. The knee is one of the most commonly injured regions of the body in sports. Following injury, it is important to be able to measure a patient's knee function to determine when they are able to return to play. Your participation in this study will be used to help evaluate and improve these instruments, thus proving better quality of care for injured athletes.

Risks/Benefits: The goal of this study is to evaluate the quality of this new measurement tool. The anticipated risk for this study are minimal. To minimize the potential risk to your privacy, your questionnaire data will be anonymized.

Confidentiality: Hard copies of the consent forms and all participant data will be stored separately within secured lockable storage cabinets within a university professor's office. Only the study investigator will have access to the data.

Principle Investigator/Contact Information: James L. Farnsworth II | Phone: (712) 749-2177 |

Email: farnsworth@bvu.edu

Participating in this project is voluntary, and refusal to participate or withdrawing from participation at any time during the project will involve no penalty or loss of benefits to which you might otherwise be entitled. All efforts, within reason, will be made to keep the personal information in your research record private but total privacy cannot be promised, for example, your information may be shared with the Buena Vista University Institutional Review Board. In the event of questions or difficulties of any kind during or following participation, you may contact the principle investigator as indicated above. For additional information about giving consent or your rights as a participant in this study, please feel free to contact the BVU IRB Chair Dr. Thom Bonagura (bonagura@bvu.edu).

Consent

I have read the above information and my questions have been answered satisfactorily by project staff. I believe I understand the purpose, benefits, and risks of the study and give my informed and free consent to be a participant.

SIGNATURE

DATE

APPENDIX F: RECRUITMENT SCRIPT CHAPTER IV

Hello student athletes,

If you have already completed our survey, THANK YOU. If not, see below and please consider completing this survey so that we can continue to improve patient-reported health outcomes for athletes.

Following injury health care providers use questionnaires to track changes in function throughout the rehabilitation process. These instruments are used to compare treatment options and ensure that patients are receiving the best quality of care. Although there are many instruments used to evaluate knee-related function following injury, none of these instruments were designed specifically for use with athletes. As a result, these instruments have limited usefulness as athletes return to normal functional levels.

Recently we developed a new instrument for measuring knee function specifically for athletes following injury. It is important to evaluate the quality of this new measure to make certain that it is accurately measuring patient function. Therefore, we are asking your assistance to help us with testing our new instrument.

We are seeking collegiate athletes who are currently free from injury to complete our survey. Please take 10-20 minutes out of your busy schedule by taking an online survey, which will ask you questions related to your knee. As compensation for your participation in this study you have the option of being entered into a drawing for one of four \$25 Visa gift cards. Submit a screen shot of the completion screen via email to humanperformance@gmail.com. If selected, you will be contacted by the primary investigator via email to obtain your mailing information for the gift card.

This research project is being conducted by Jim Farnsworth from the School of Education and Exercise Science at Buena Vista University.

There are no known risks if you decide to participate in this research study, nor are there any costs for your participation. This survey is anonymous. There is no way to ensure full anonymity using the internet; however, we will be using an encrypted online survey tool. No one will be able to identify you, nor will anyone be able to determine your place of employment. No one will know whether you participated in this study.

Your participation in this study is voluntary. If you choose to participate, please follow the link: <https://tinyurl.com/knee-function2> and proceed to the survey.

If you have any questions about the study, please contact professor Jim Farnsworth (712) 249-2177 or by email farnsworth@bvuv.edu.

If you have any questions about your rights as a research subject, you may contact the Buena Vista University Institutional Review Board (IRB) by email bonagura@bvuv.edu.

Thank you,
Jim Farnsworth

APPENDIX G: RECRUITMENT FLYER CHAPTER IV



Seeking Athletes for Survey Research!

Purpose: The purpose of this study is to evaluate injury outcome measures in athletics. Many of the current instruments don't do a good job of tracking athlete recovery after an injury. We are seeking *athletes to help* us by completing a brief survey online.

We will be conducting a raffle at the end of the study to give away one of four \$25 Visa gift cards. The survey can be completed in person as well. For more details about the study please contact Jim Farnsworth or visit the link below.

Link to Online Survey: The link to the survey for this study can be found online at:
<http://tinyurl.com/PRKOATv2>

Contact Information: **Office:** (712) 749-2177 **Email:** farnsworth@bv.edu



APPENDIX H: PRKOAT (Pilot Version)

Today's date:

Date of birth:

____/____/____

____/____/____

Patient Full Name:

INSTRUCTIONS: The following questionnaire will ask you questions about your knee function. This information will be used to help us evaluate your knee symptoms, pain, and how well you are able to perform usually daily activities. Read each question carefully and circle only **one** answer per question. Some of the items may not be directly relevant to your specific condition. If you are unsure about how to answer a question, please give the best answer that you can.

1) How severe is your knee joint stiffness after first waking in the morning?

No Stiffness

Mild

Moderate

Severe

2) How severe is your knee joint stiffness after sitting, lying or resting later in the day?

No Stiffness

Mild

Moderate

Severe

3) How often do you experience knee pain?

Never

Rarely

Sometimes

Almost Always

What amount of knee pain have you experienced the **last week** during the following activities?

4) Pain with twisting/pivoting on your knee?

None

Mild

Moderate

Extreme

5) Pain with straightening knee fully?

Never

Rarely

Sometimes

Almost Always

6) Pain with bending knee fully?

Never

Rarely

Sometimes

Almost Always

7) Pain with walking on a flat surface?

Never

Rarely

Sometimes

Almost Always

8) Pain with going up or down stairs?

Never

Rarely

Sometimes

Almost Always

9) Pain at night while in bed?

Never

Rarely

Sometimes

Almost Always

10) Pain with sitting or lying?

Never

Rarely

Sometimes

Almost Always

For each of the following activities please indicate the degree of difficulty you have experienced in the **last week** due to your knee.

11) Difficulty with ascending/descending stairs?

None

Mild

Moderate

Extreme

12) Difficulty with rising from sitting?

None

Mild

Moderate

Extreme

13) Difficulty with getting in/out of your vehicle?

None

Mild

Moderate

Extreme

14) Difficulty with standing?

None

Mild

Moderate

Extreme

15) Difficulty with bending to the floor?

None

Mild

Moderate

Extreme

16) Difficulty with picking up a small object (i.e., penny) off the floor?

None	Mild	Moderate	Extreme
------	------	----------	---------

17) Difficulty with lifting a large object off the floor?

None	Mild	Moderate	Extreme
------	------	----------	---------

18) Difficulty with walking on a flat surface?

None	Mild	Moderate	Extreme
------	------	----------	---------

19) Difficulty with sitting with your knee bent?

None	Mild	Moderate	Extreme
------	------	----------	---------

20) Difficulty with performing double leg squats (body weight only)?

None	Mild	Moderate	Extreme
------	------	----------	---------

21) Difficulty with performing a single leg squat on your injured knee?

None	Mild	Moderate	Extreme
------	------	----------	---------

22) Difficulty with lower body resistance training (e.g., weighted lunges, weighted squats, etc.)

None	Mild	Moderate	Extreme
------	------	----------	---------

23) Difficulty with jogging (approximately 50% intensity)?

None	Mild	Moderate	Extreme
------	------	----------	---------

24) Difficulty with running at maximum speed?

None	Mild	Moderate	Extreme
------	------	----------	---------

25) Difficulty with quickly changing direction while running (e.g., agility drills)?

None	Mild	Moderate	Extreme
------	------	----------	---------

26) Difficulty with jumping forward and landing on your injured knee?

None	Mild	Moderate	Extreme
------	------	----------	---------

27) Difficulty with jumping to the side and landing on your injured knee?

None	Mild	Moderate	Extreme
------	------	----------	---------

28) Difficulty with hopping repeatedly on your injured knee?

None	Mild	Moderate	Extreme
------	------	----------	---------

29) Difficulty with twisting/pivoting on your injured knee?

None	Mild	Moderate	Extreme
------	------	----------	---------

30) Difficulty with kneeling on the front of your knee?

None

Mild

Moderate

Extreme

31) How often are you aware of your knee problems?

Never

Rarely

Sometimes

Almost Always

32) Have you modified your lifestyle to avoid potentially damaging activities?

Never

Rarely

Sometimes

Almost Always

33) How much are you troubled with lack of confidence in your knee?

Not at all

Mildly

Moderately

Totally

34) What is the highest level of activity that you can perform without significant knee pain?

Strenuous
activities like
jumping or
pivoting as in
basketball or
soccer

Moderate
activities like
running or jogging

Light activities like
walking,
housework or
yardwork

Unable to perform
any of the above
activities due to
knee pain

35) What is the highest level of activity that you can perform without significant swelling in your knee?

Strenuous	Moderate	Light activities like	Unable to perform
activities like	activities like	walking,	any of the above
jumping or	running or jogging	housework or	activities due to
pivoting as in		yardwork	knee pain
basketball or			
soccer			

36) What is the highest level of activity that you can perform without significant giving way in your knee?

Strenuous	Moderate	Light activities like	Unable to perform
activities like	activities like	walking,	any of the above
jumping or	running or jogging	housework or	activities due to
pivoting as in		yardwork	knee pain
basketball or			
soccer			

37) Compared to your non-injured peers (e.g., teammates or friends of similar athletic ability) how would you rate your overall level of knee function?

Worse	About the same	Better
-------	----------------	--------

38) Compared to your healthy knee how would you rate your injured knee's overall level of function?

Worse

About the same

Better

39) Compared to your non-injured peers (e.g., teammates or friends of similar athletic ability) how would you rate your usual level of physical activity?

Worse

About the same

Better

40) In the box below please list the five most difficult sports-related activities, involving your knee, that you perform on a regular basis. Examples of activities may include (but are not limited to): stopping while running, jumping, repetitive movement, rising from a prone (face down) or supine (face up) position.

--

41) Rank the difficulty of the 5 activities you listed in the previous question with 5 being the most difficulty and 1 being the least difficult.

42) Are there any questions related to your knee that you feel should have been asked in this questionnaire that were not included? If all questions were answered satisfactorily then write none in the box below.

***Thank you very much for completing all of the questions in this
questionnaire.***

APPENDIX I: PRKOAT (Revised)**PATIENT REPORTED KNEE OUTCOMES ASSESSMENT TOOL****(PRKOAT)**

Today's date:

Date of birth:

____/____/____

____/____/____

Patient Full Name:

INSTRUCTIONS: The following assessment will ask you questions about your knee function. This information will be used to help us evaluate your knee symptoms, pain, and how well you are able to perform usual daily activities. Read each question carefully and circle only **one** answer per question. If you are unsure about how to answer a question, please give the best answer that you can.

1) How severe is your knee joint stiffness after first waking in the morning?

No Stiffness

Mild

Moderate

Severe

2) How severe is your knee joint stiffness after sitting, lying or resting later in the day?

No Stiffness

Mild

Moderate

Severe

3) How often do you experience knee pain?

Never Rarely Sometimes Almost Always

What amount of knee pain have you experienced in the **last week** during the following activities?

4) Pain while twisting/pivoting on your knee?

Never Rarely Sometimes Almost Always

5) Pain while straightening knee fully?

Never Rarely Sometimes Almost Always

6) Pain while bending knee fully?

Never Rarely Sometimes Almost Always

7) Pain while going up or down stairs?

Never Rarely Sometimes Almost Always

8) Pain while sitting or lying?

Never Rarely Sometimes Almost Always

9) Pain while bending to the floor and picking up a small object (such as a penny) off of the floor?

Never	Rarely	Sometimes	Almost Always
-------	--------	-----------	---------------

10) Pain while standing for an extended period of time (minimum 1 hour)?

Never	Rarely	Sometimes	Almost Always
-------	--------	-----------	---------------

11) Pain while sitting with your knee bent (at any degree) for an extended period of time (minimum 1 hour)?

Never	Rarely	Sometimes	Almost Always
-------	--------	-----------	---------------

12) Pain while rising from an armless chair?

Never	Rarely	Sometimes	Almost Always
-------	--------	-----------	---------------

13) Pain while lifting a large object off of the floor?

Never	Rarely	Sometimes	Almost Always
-------	--------	-----------	---------------

14) Pain while performing standard squats (body weight only)?

Never	Rarely	Sometimes	Almost Always
-------	--------	-----------	---------------

15) Pain while performing single leg squats on your injured leg (body weight only)?

Never	Rarely	Sometimes	Almost Always
-------	--------	-----------	---------------

16) Pain while performing lower-body resistance training exercises with weight (such as lunges, step-ups, or squats)?

Never	Rarely	Sometimes	Almost Always
-------	--------	-----------	---------------

17) Pain while jogging at approximately 50% intensity?

Never	Rarely	Sometimes	Almost Always
-------	--------	-----------	---------------

18) Pain while running at maximum speed?

Never	Rarely	Sometimes	Almost Always
-------	--------	-----------	---------------

19) Pain while back pedaling (i.e., running backwards)?

Never	Rarely	Sometimes	Almost Always
-------	--------	-----------	---------------

20) Pain associated with changing direction while running (e.g., agility drills)?

Never	Rarely	Sometimes	Almost Always
-------	--------	-----------	---------------

21) Pain with jumping forward and landing on your injured knee?

Never	Rarely	Sometimes	Almost Always
-------	--------	-----------	---------------

22) Pain with jumping to the side and landing on your injured knee?

Never	Rarely	Sometimes	Almost Always
-------	--------	-----------	---------------

23) Pain with repeated high-impact jumps in place (such as jump knee tucks)?

Never	Rarely	Sometimes	Almost Always
-------	--------	-----------	---------------

24) Pain with kneeling on your injured knee?

Never	Rarely	Sometimes	Almost Always
-------	--------	-----------	---------------

25) How often are you aware of your knee problems?

Never	Rarely	Sometimes	Almost Always
-------	--------	-----------	---------------

26) Have you modified your daily exercise or practice routines to avoid painful or potentially damaging activities?

Never	Rarely	Sometimes	Almost Always
-------	--------	-----------	---------------

27) To what extent are you troubled with lack of confidence in your knee when engaging in intense physical activity (such as during practice or games)?

Not at all Mildly Moderately Totally

28) To what extent do you feel anxious about performing certain activities because of your injured knee?

Not at all Mildly Moderately Totally

29) What is the highest level of activity that you can perform without significant swelling in your knee?

Strenuous	Moderate	Light activities like	Unable to
activities like	activities like	walking to class,	perform any of
jumping or	running or	cleaning your	the above
pivoting as in	jogging	apartment/dorm	activities due to
basketball or		room	knee pain
soccer			

30) Compared to your non-injured peers (e.g., teammates or friends of similar athletic ability) how would you rate your overall level of knee function?

Worse About the same Better

31) Compared to your healthy knee how would you rate your injured knee's overall level of function?

Worse

About the same

Better

32) Compared to your non-injured peers (e.g., teammates or friends of similar athletic ability) how would you rate your usual level of physical activity?

Worse

About the same

Better

Thank you very much for completing all of the questions in this questionnaire.